

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:39 ; Search time 100.46 Seconds

(without alignments)
1384.761 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398

Sequence: 1 MRYRASALGSDGVRTMESA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SID82/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SID82/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SID82/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SID82/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5216	96.6	1063	AA039991	Human polypeptide
2	5165.5	95.7	1050	ABG10171	Novel human diapo
3	3369	62.4	935	ABG10172	Novel human diapo
4	2378	44.1	537	AA041777	Human polypeptide
5	1800.5	33.0	1011	AB071113	Drosophila melanog
6	1782	33.4	700	AA064615	Human secreted pro
7	1770.5	32.8	697	AA064616	Human secreted pro
8	1738.5	32.2	1026	AA073125	Novel human enzyme
9	1714.5	31.8	1098	AA075544	Human Myosin-1F pr
10	1685	31.2	909	AA023128	Novel human enzyme

11	1565	29.0	1089	AB097258	Novel human protei
12	1412.5	26.2	759	AB020610	Novel human diapo
13	1362	25.2	2167	AB060369	Drosophila melanog
14	1330.5	24.6	512	AB02651	Human ORF2415
15	1328	24.6	2129	AB062828	Drosophila melanog
16	1301.5	24.1	1493	AB048639	Arabidopsis thalia
17	1301.5	24.1	1544	AB048639	Arabidopsis thalia
18	1294	24.0	1483	AB048638	Arabidopsis thalia
19	1284.5	23.8	2058	AB048640	Arabidopsis thalia
20	1272.5	23.6	1502	AB097219	Novel human protei
21	1269	23.5	1504	AB058924	Drosophila melanog
22	1269	23.5	1556	AB030439	Arabidopsis thalia
23	1261.5	23.4	1495	AB030440	Arabidopsis thalia
24	1241.5	23.0	2048	AA011891	Arabidopsis thalia
25	1241.5	23.0	2057	AA011890	Arabidopsis thalia
26	1223.5	22.7	1754	AA052410	Arabidopsis thalia
27	1223.5	22.7	1769	AA052409	Arabidopsis thalia
28	1223.5	22.7	1804	AA052408	Arabidopsis thalia
29	1218	22.6	1839	AA054291	Caenorhabditis ele
30	1216	22.5	764	AA080123	Human protein SEQ
31	1208	22.4	697	AA000603	Human non-conventi
32	1191	22.1	753	AA079139	Human protein SEQ
33	1191	22.1	1939	AB077096	Human alpha-myosin
34	1187.5	22.0	1453	AA039213	Human polypeptide
35	1186	22.0	369	AA057005	Human prostate can
36	1183.5	21.9	1469	AA039214	Human polypeptide
37	1182	21.9	1120	AA094292	Helianthus annuus
38	1181.5	21.9	631	AA000604	Murine myosin VII
39	1181	21.9	1879	AA025750	Human protein sequ
40	1174.5	21.8	1960	AA078854	Human protein SEQ
41	1174.5	21.8	1963	AA079838	Human protein SEQ
42	1172.5	21.7	2056	AB059344	Drosophila melanog
43	1171	21.7	1988	AA040999	Human polypeptide
44	1171	21.7	1988	AA041000	Human polypeptide
45	1169.5	21.7	1972	AA000024	Smooth muscle myos

ALIGNMENTS

RESULT 1	AA039991	standard; Protein; 1063 AA.
ID	AA039991	
XX	AA039991	
AC	AA039991	
XX	22-OCT-2001 (first entry)	
XX	Human polypeptide SFQ ID NO 3136.	
XX		
DE	Human: noctropic; immunosuppressant; cytostatic; gene therapy; cancer;	
XX	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukemia.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153312-A1.	
PD	26-JUL-2001.	
XX		
XX	26-DEC-2000; 2000WO-US34263.	
PF		
XX	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	
PR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	

XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
 DR WPI: 2001-442253/47.
 DR N-PSDB: AA159147.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Example 4; SEQ ID NO 3136; 10078bp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 XX the encoded polypeptides (AA158642-AA162213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 XX
 SQ Sequence 1063 AA:
 Query Match 96.6%; Score 5216; DB 22; Length 1063;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1003; Conservative 22; Mismatches 13; Indels 0; Gaps 0;
 QY 7 AAGSGVRYTMSALTARVGVODVLENTSEAFENLRFRRENTLYTYIGPVLY 66
 DB 26 AAGSGVRYTMSALTARVGVODVLENTSEAFENLRFRRENTLYTYIGPVLY 85
 QY 67 SVNPRDLOIYSRQHMERIRGVSEYEPHLEFAVDYTRALRTERDQAVMISSESGAG 126
 DB 86 SVNPRDLOIYSRQHMERIRGVSEYEPHLEFAVDYTRALRTERDQAVMISSESGAG 145
 QY 127 KTEATRLQFAETCPAREGAVDRLOSNPYLEAFGNKTLRNDSSRFKYMDOV 186
 DB 146 KTEATRLQFAETCPAREGAVDRLOSNPYLEAFGNKTLRNDSSRFKYMDOV 205
 QY 187 FDEKAPVGHILSYLLEKSRVYHONHGERNFHVFYQLLEGESEETLRRLGLEPNOSTYL 246
 DB 206 FDEKAPVGHILSYLLEKSRVYHONHGERNFHVFYQLLEGESEETLRRLGLEPNOSTYL 265
 QY 247 YLVKGCACVSSINDSKMKVARKALSVDFTEDEVEDLSTVASYLHIGNHFADEDS 306
 DB 266 YLVKGCACVSSINDSKMKVARKALSVDFTEDEVEDLSTVASYLHIGNHFADEDS 325
 QY 307 NAGVTENOLKYLTRLGVEGTLRREALTHRKIIANGEBELSPNLNQAAYARDALAKAV 366
 DB 326 NAGVTENOLKYLTRLGVEGTLRREALTHRKIIANGEBELSPNLNQAAYARDALAKAV 385
 QY 367 YSRFTWLVKIRNSLASKDAESPWSRSTTVGLLDIYGEVFOHNSFQFCINTYNEKL 426
 DB 386 YSRFTWLVKIRNSLASKDAESPWSRSTTVGLLDIYGEVFOHNSFQFCINTYNEKL 445
 QY 427 QQLFIETLTKSEDEEAEAGIAMEPOYFNKKITCDLVEKFKGIISIDDECLRGCEAT 486
 DB 446 QQLFIETLTKSEDEEAEAGIAMEPOYFNKKITCDLVEKFKGIISIDDECLRGCEAT 505
 QY 487 DLTFLEKLEDTVRPHFELTHKLADQTRKSLDRGEFRLIHYAGEVTVSVTGLDRKNDL 546
 DB 506 DLTFLEKLEDTVRPHFELTHKLADQTRKSLDRGEFRLIHYAGEVTVSVTGLDRKNDL 565

QY 547 LERNKETMSSNPIAOCFDPKSELSDKRRETVATOFKMSILOLVEILRSKPEAYIRC 606
 DB 566 LERNKETMSSNPIAOCFDPKSELSDKRRETVATOFKMSILOLVEILRSKPEAYIRC 625
 QY 607 IKPNDAKOPGRDEVLIRHQQVYLGIMENLRYRRAGFARRYEAFLOQRYSICPEWPM 666
 DB 626 IKPNDAKOPGRDEVLIRHQQVYLGIMENLRYRRAGFARRYEAFLOQRYSICPEWPM 685
 QY 667 WAGRPODGAVALVRHLGTYPEEKMGRTKIFIRFETPATDSLEVRQSLATRIQAAW 726
 DB 686 WAGRPODGAVALVRHLGTYPEEKMGRTKIFIRFETPATDSLEVRQSLATRIQAAW 745
 QY 727 RGFHMRKELRYRSACIQSMWRGTLGRRAKAKKAAQTIRRLIRGILHNSRCPEN 786
 DB 746 RGFHMRKELRYRSACIQSMWRGTLGRRAKAKKAAQTIRRLIRGILHNSRCPEN 805
 QY 787 AFLDHYRASFLNLRQLPRLNLTSPPLPALREASELLRELCKMMWKYCISISP 846
 DB 806 AFLDHYRASFLNLRQLPRLNLTSPPLPALREASELLRELCKMMWKYCISISP 865
 QY 847 EMKQLOQKAVASEIFKGGKDNTPQSVPLFISTRLGTEISPRVLOSISEPIQYAVV 906
 DB 866 EMKQLOQKAVASEIFKGGKDNTPQSVPLFISTRLGTEISPRVLOSISEPIQYAVV 925
 QY 907 VKYDRKGYRPPQOLITPSAVYVEDAKYKORIDVANLIGISVSLSDIFLHVQRD 966
 DB 926 VKYDRKGYRPPQOLITPSAVYVEDAKYKORIDVANLIGISVSLSDIFLHVQRD 985
 QY 967 NKQGDVVLSDHYETLFTALSDRVNNININOSITPAGGPGHDIIFTSSELLI 1026
 DB 986 NKQGDVVLSDHYETLFTALSDRVNNININOSITPAGGPGHDIIFTSSELLI 1045
 QY 1027 TKAKNGHLAVVAPRLNSR 1044
 DB 1046 TKAKNGHLAVVAPRLNSR 1063
 RESULT 2
 ABG10171
 ID ABG10171 standard; Protein; 1050 AA.
 XX 'ABG10171;
 XX 13-FEB-2002 (first entry)
 DE Novel human diagnostic protein #10162.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 OS food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 PN WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-0508631.
 PE 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB: AA574358.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnosis, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

xx Claim 20: SEQ ID No 40530: 103bp: English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 xx Sequence 1050 AA:
 Query Match 95.7%; Score 5165.5; DB 22; Length 1050;
 Best Local Similarity 95.8%; Pred. No. 0;
 Matches 995; Conservative 24; Mismatches 19; Indels 1; Gaps 1;
 QY 7 ALGSDGVATMESALTARDRVGVDFVLENTSFAFIENLRFRRENLIYTYIGPYLV 66
 DB 12 ALGSDGVATMESALTARDRVGVDFVLENTSFAFIENLRFRRENLIYTYIGPYLV 71
 QY 67 SVNPRDLQIYSRQHMERYRGVSFEYVPHLEFAVDYVYRALRTERRDQAVMISGESGAG 126
 DB 72 SVNPRDLQIYSRQHMERYRGVSFEYVPHLEFAVDYVYRALRTERRDQAVMISGESGAG 131
 QY 127 KTEARKRLQFYAEICPAERPGAVRDRLQSNPYLAEFGNAKTLRNNSRFRGKYMDVQ 186
 DB 132 KTEARKRLQFYAEICPAERPGAVRDRLQSNPYLAEFGNAKTLRNNSRFRGKYMDVQ 191
 QY 187 FDFKAPYVGHILSYLLEKSRVYHONHGERNFHFVYQLLEGSEETLRRLLENNPOSYL 246
 DB 192 FDFKAPYVGHILSYLLEKSRVYHONHGERNFHFVYQLLEGSEETLRRLLENNPOSYL 251
 QY 247 YLVKQCAKAVSSINDSKDMKYRKALSYIDFTEDEVEDLLSIVASVHLGNHFAADEDS 306
 DB 252 YLVKQCAKAVSSINDSKDMKYRKALSYIDFTEDEVEDLLSIVASVHLGNHFAADEDS 311
 QY 307 NAOVVTENQKYLTPFLLEVEGTLTREALTHRKIIAKGSELLSPNLDEQAARAADLAAY 366
 DB 312 NAOVVTENQKYLTPFLLEVEGTLTREALTHRKIIAKGSELLSPNLDEQAARAADLAAY 371
 QY 367 YSRFTFWLVRKINRSLASKDAESPSMRSTVYGLLDIYGFVFOHNSFOFCINYNCKEL 426
 DB 372 YSRFTFWLVRKINRSLASKDAESPSMRSTVYGLLDIYGFVFOHNSFOFCINYNCKEL 431
 QY 427 QQLFTLTLKSEQBEYEAGIAMEPVQYFNKKIICDLVEKFKGIISILDEECLRGSEAT 486
 DB 432 QQLFTLTLKSEQBEYEAGIAMEPVQYFNKKIICDLVEKFKGIISILDEECLRGSEAT 491
 QY 487 DLTFLKLELDYKPRPHFLTKHLAQKTRKSLDRREFLLHYAGVTVSYVGLDKNDL 546
 DB 492 DLTFLKLELDYKPRPHFLTKHLAQKTRKSLDRREFLLHYAGVTVSYVGLDKNDL 551
 QY 547 LFRNLKETWCSMNPIMAQCFDSELSDKKRPETVATQFKMSLDLVEILSKSEPAYTRC 606
 DB 552 LFRNLKETWCSMNPIMAQCFDSELSDKKRPETVATQFKMSLDLVEILSKSEPAYTRC 611
 QY 607 IKPNDAKQGRFDEVILRHQVYKYLGLMENLVRRAAGFAYRRKYEAFLORYKSLCPETWPM 666

DB 612 IKPNDAKQGRFDEVILRHQVYKYLGLMENLVRRAAGFAYRRKYEAFLORYKSLCPETWPM 671
 QY 667 WAGRPODGAVALVRHLGYKPEEYKMGRTKIFRPEKTLFATEDSLVRROSLATKIOAM 726
 DB 672 WAGRPODGAVALVRHLGYKPEEYKMGRTKIFRPEKTLFATEDSLVRROSLATKIOAM 731
 QY 727 RGFHMROKFLVRVRSACIOSMWRGTLGRKAKRKMAAQTIRLIRGFLRHSRCPEN 786
 DB 732 RGFHMROKFLVRVRSACIOSMWRGTLGRKAKRKMAAQTIRLIRGFLRHSRCPEN 791
 QY 787 AFFLDVHRSFLLNLRRLQRLPRNLDTSWTPPALREASBELRELCKMMWVKYCRSIS 846
 DB 792 AFFLDVHRSFLLNLRRLQRLPRNLDTSWTPPALREASBELRELCKMMWVKYCRSIS 851
 QY 847 EMKROOQKRVASEIRKGGKDNVPOSVPRFISTRGTEISRVLQSIGSEPIQAVPV 906
 DB 852 EMKROOQKRVASEIRKGGKDNVPOSVPRFISTRGTEISRVLQSIGSEPIQAVPV 911
 QY 907 VKYDRKGYKRPRLQLLTPS-AVVIYEDAKVKORIDYANLTGISVSLDSLFVLYHORE 965
 DB 912 VKYDRKGYKRPRLQLLTPS-AVVIYEDAKVKORIDYANLTGISVSLDSLFVLYHORE 971
 QY 966 DNKQKGVYVQSDHYETLTKTALSADRVNNININGSTIFAGGPRDGIIDTSSSELL 1025
 DB 972 DNKQKGVYVQSDHYETLTKTALSADRVNNININGSTIFAGGPRDGIIDTSSSELL 1031
 QY 1026 ITRAKNGHLAVVAPRLNSR 1044
 DB 1032 ITRAKNGHLAVVAPRLNSR 1050
 RESULT 3
 ABG10172
 ID ABG10172 standard; Protein; 935 AA.
 XX
 AC ABG10172;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #10163.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 XX N-PSDB: NAA574359.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits and to assess
 XX biodiversity -
 PS Claim 20: SEQ ID No 40531; 103bp: English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (i) is useful in gene therapy techniques
 CC to restore normal activity of (ii) or to treat disease states involving
 CC (ii). (ii) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (ii) and its binding partners are useful in medical
 CC imaging of sites expressing (ii). (i) and (ii) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 935 AA;

Query Match 62.4%; Score 3369; DB 22; Length 935;
 Best Local Similarity 86.0%; Pred. No. 6, 2e-278;
 Matches 662; Conservative 22; Mismatches 20; Indels 66; Gaps 6;

QY 190 KGAPVGHILSYLLEKSRVYVHNGERNFHYQLLEGGEETLRRLGLERNPQSYLYLV 249
 DB 82 KGAPVGHILSYLLEKSRVYVHNGERNFHYQLLEGGEETLRRLGLERNPQSYLYLV 141
 QY 250 KGCAVYSINDSKDMVKAKALSYDFTFDEVEDLLSYASVHLGNHFADEDSNAQ 309
 DB 142 KGCAVYSINDSKDMVKAKALSYDFTFDEVEDLLSYASVHLGNHFADEDSNAQ 201
 QY 310 VTTENOLKYLTRLVEGTTLRRLTHRIKIIAGEFELSPLENOAAYARDALAKVYSR 369
 DB 202 VTTENOLKYLTRLVEGTTLRRLTHRIKIIAGEFELSPLENOAAYARDALAKVYSR 261
 QY 370 TFWVLVAKINRSLSK--DAESPWR--STTVGLDLIDYGEVDFHNSFEQFCINYN 423
 DB 262 TFWVLVAKINRSLSK--DAESPWR--STTVGLDLIDYGEVDFHNSFEQFCINYN 313
 QY 424 EKLOOLFIELTLEKSEYEAEIGIAMEPVOYFNKTIICLVKEKGIISIDEELRAG 483
 DB 314 EKLOOLFIELTLEKSEYEAEIGIAMEPVOYFNKTIICLVKEKGIISIDEELRAG 373
 QY 484 EATDLTLEKLEDTVRPHFHTLTKLADQTRKSLDRGEFRLIHYAGEVYSYTGFLDN 543
 DB 374 EATDLTLEKLEDTVRPHFHTLTKLADQTRKSLDRGEFRLIHYAGEVYSYTGFLDN 433
 QY 544 NDLLFRLNLEKSSNMPIMAGCFDSELSDKRPPTVAFORFMSLLQVLEILRSKPAY 603
 DB 434 NDLLFRLNLEKSSNMPIMAGCFDSELSDKRPPTVAFORFMSLLQVLEILRSKPAY 493
 QY 604 IRCIRPNDAKO-----GGRFDEVIRHOVYVGLIEMENLVRRRGEFAR 646
 DB 494 IRCIRPNDAKO-----GGRFDEVIRHOVYVGLIEMENLVRRRGEFAR 541
 QY 647 RKYEAFLOKYSKLPETPMWAGRPDGYAVLVRHLGYRPEEKMGRTIFTRPKTLFA 706
 DB 542 RKYEAFLOKYSKLPETPMWAGRPDGYAVLVRHLGYRPEEKMGRTIFTRPKTLFA 601
 QY 707 TEDSELEVRQSLATKIQAAWGRFHMROKRLRYKRSALICQSWMGRTGLGRKAKRMAQ 766
 DB 602 TEDSELEVRQSLATKIQAAWGRFHMROKRLRYKRSALICQSWMGRTGLGRKAKRMAQ 638
 QY 767 TIRRLIRGFTLRHSPRCENAFELDHVRASEFLNLRQPRVAVLDTSWTPPALRESE 826
 DB 639 TIRRLIRGFTLRHSPRCENAFELDHVRASEFLNLRQPRVAVLDTSWTPPALRESE 698
 QY 827 LIRELCMKNMWKCYSISPEWKOLOQKAVASEIFKGRKNDYPOSPVPLFTSTRIGTEE 886
 DB 699 LIRELCMKNMWKCYSISPEWKOLOQKAVASEIFKGRKNDYPOSPVPLFTSTRIGTEE 758
 QY 887 ISPRVLOLGSSEPIOTAVPVVYKDRKGKPRRQLLTPSAVIVEDAKV 936

DB 759 ISPRVLOLGSSEPIOTAVPVVYKDRKGKPRRQLLTPSAVIVEDAKV 808

RESULT 4
 AAM41777
 ID AAM41777 standard; Protein; 537 AA.

AC AAM41777;

XX 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6708.

XX Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI: 2001-442253/47.

PS N-PSDB: AAI60933.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

PS Example 2; SEQ ID NO 6708; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nocotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 537 AA;

Query Match 44.1%; Score 2378; DB 22; Length 537;
 Best Local Similarity 96.7%; Pred. No. 9, 1e-194;

Matches 462; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 7 ALGSGVATWESALTARDRVGVDFVLENTSEAAFIENLRFRFENLTYTIGVIV 66
 DB 60 ALGSGVATWESALTARDRVGVDFVLENTSEAAFIENLRFRFENLTYTIGVIV 119
 QY 67 SVNPYRDQIYSRQIMERYRGVSFEVPPHFAVADYVYALRTERRDQAVMISGESGAG 126
 DB 120 SVNPYRDQIYSRQIMERYRGVSFEVPPHFAVADYVYALRTERRDQAVMISGESGAG 179
 QY 127 KTEATKRLLOFYAETCPAPERGAVRDRLQSNPYLEAFNAKTLRNDSSRFEGKYMDVQ 186
 DB 180 KTDATKRLLOFYAETCPAPERGAVRDRLQSNPYLEAFNAKTLRNDSSRFEGKYMDVQ 239
 QY 187 FDFKAPVGHILSLLEKSRVYVHNGERNPHVYQULLEGGEETLRRLGLERNPOSYL 246
 DB 240 FDFKAPVGHILSLLEKSRVYVHNGERNPHVYQULLEGGEETLRRLGLERNPOSYL 299
 QY 247 YLVKGCARVSSINDKSDMKVYRKALSVIDTEDEVEDLLSTVAVLHGINHFAADEDS 306
 DB 300 YLVKGCARVSSINDKSDMKVYRKALSVIDTEDEVEDLLSTVAVLHGINHFAADEDS 359
 QY 307 NAOVTTENOLKYLRLGVEGTTLRREALTHRKIIAKGELLSPNLBQAAVARDALAKAV 366
 DB 360 NAOVTTENOLKYLRLGVEGTTLRREALTHRKIIAKGELLSPNLBQAAVARDALAKAV 419
 QY 367 YSRFTTWLVKRNISLAKDAESPMSRSTVYGLDIYGFYFQHNSEFOCINCNKL 426
 DB 420 YSRFTTWLVKRNISLAKDAESPMSRSTVYGLDIYGFYFQHNSEFOCINCNKL 479
 QY 427 QOLFELTLKSEOEYEAEIGIAMEPVQYFNKKIICDLVBEKFKGIISLDEBCLRGE 484
 DB 480 QOLFELTLKSEOEYEAEIGIAMEPVQYFNKKIICDLVBEKFKGIISLDEBCLRGE 537

RESULT 5
 ABB71113
 ID ABB71113 standard; Protein; 1011 AA.
 AC ABB71113;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 40131.
 XX Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EM;
 PI MPI: 2001-656860/75.
 DR N-PSDB; ABL15216.
 XX New Isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 40131; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB157737-AB172072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1011 AA;
 SQ

Query Match 33.4%; Score 1800.5; DB 22; Length 1011;
 Best Local Similarity 39.6%; Pred. No. 4,5e-144;
 Matches 426; Conservative 169; Mismatches 341; Indels 139; Gaps 29;

QY 28 GVODFVLENTSEAAFIENLRFRFENLTYTIGVIVSVNPYRDQIYSRQIMERYRG 87
 DB 8 GVODFVLLDQVSMH-KEMDMILRRFONGSIYTYIGEVCVSMNPYRQMINIYGPETIRKYK 66
 QY 88 VSFYEVPPHFAVADYVYALRTERRDQAVMISGESGAGKTEATKRLLOFYAETCPAP-- 145
 DB 67 RELFENAPHLFALDASVRYLAKORQDTCILISGESGAGKTEATKIIAAVTNAGQ 126
 QY 146 ---ERGAVRDRLQSNPYLEAFNAKTLRNDSSRFEGKYMDVQDFKAPVGHILSYL 202
 DB 127 NEIR---VKNVLIQASNAILETQGNATKNNDSSRFEGKMDLEDFDKAPVGGIITNVL 183
 QY 203 LEKSRVYVHNGERNPHVYQULLEGGEETLRRLGLERNPOSYLVLVKGCAVSSINDK 262
 DB 184 LEKSRVYVQOQGEENFHSFYOLLRGANDNEIROYELQKFGKHYHLOG---SMDILTEK 240
 QY 263 SDMKVYRKALSVIDTEDEVEDLLSTVAVLHGINHFAADEDSNAOVTTENOLKYLTRL 322
 DB 241 SDYKGTNAEFTLGFSTDEVOITWRTIAAVLHAGNVEFOTIED-ELVYSKQHLKSTAKL 299
 QY 323 LGVEGTTLRREALTHRKIIANGELLSPNLBQAAVARDALAKAVYSTFTWLVKRNISL 382
 DB 300 LQVTELTSTALTRVYLAAGSNVQKDNMTQAVRGDALAKAIYDRLFTWISIRNRAI 359
 QY 383 ASKDAESPSMRSTVYGLDIYGFYFQHNSEFOCINCNKIQOLFELTLKSEOEY 442
 DB 360 LFRSQTQA-RFNSVIGLIDYGFELPDSNFEFOCINCNKIQOLFELTLKSEOEY 418
 QY 443 EAEIGIAMEPVQYFNKKIICDLVBEKFKGIISLDEBCLRGEATDLFLEKLEDTVPHP 502
 DB 419 QREGIEMNTIYFNKKIICDLVBEKFKGIISLDEBCLRGEATDLFLEKLEDTVPHP 478
 QY 503 HFLTHKLADOKTRSL-DRGEFRLHYAGEVTVSVTFLDKNDLFR-----NIKET 554
 DB 479 HYSRQL--KPTDKELKHREDFRITHYAGDVYINNGFIEKNKDTLYQDKRLHNSKDA 536
 QY 555 MCSMNPIMACQFQKSELSDKKRPYATOKKMSLDLVLETRSKEPAYTRCIRPNQAK 614
 DB 537 NLESMMPGADIRKT---TKRPLTAGTLEQRMADLVYTLKKEPPYVRCIKPNLKS 592
 QY 615 PGRFDEVYLRHGVYGLMNTLRVYRAGFAVRRYEAFLORYSICETPMW-AGRPD 673
 DB 593 STVDEDEVEHVOVYTLGLETLRRRAGVYRQYDFLRYKMIISQYTNPFAGSDRD 652
 QY 734 KFLVYKRAICIGSMWRTGLGRRKAARKMAQOTIRLRIGETLRHSPCEANFPLDHV 793
 DB 712 NFKKKK-----NATIVRYKAYKLR----- 732
 QY 794 RASFLNLRQL--PRNVL---TSMTPPALREASSELLR---ELCKMNV-WKYCRS 843
 DB 733 ---SYVQELANLRRAKQKRDYKGIOWQPPPLAKRYEAKLRHMFDFBRANMILHKYPRS 790

Query Match 32.2% Score 1738.5; DB 22; Length 1026;
 Best Local Similarity 41.6%; Pred. No. 9e-139;
 Matches 404; Conservative 155; Mismatches 339; Indels 73; Gaps 22;

```

QY 28 GVDDFVLENTSEAFIENLRRENTLITYIGVLYSVNPRDQIYSROHMERRG 87
DB 18 GKDFEVLLOQVME-DEMRNLOLREFEGRITYIGVLYSVNPEOLPLXGPEALARYG 76
QY 88 VSEYEPHLEFVAADVYRALRTERDQAVMTSGESGACKTATKRLLOFYMETCAPER 147
DB 77 RELTERPHLYAVANAYKAKMRSDTCTIVISGESGACKTAKHIMQYIAVTPSOR 136
QY 148 GGA--VRRLLOSVPYLEAFGNAKTLRNDSSRFKGYDOPDFKAPGHIISYLEK 205
DB 137 AEYERKDVLLKSTCYLEAFGNARTNRHNSRFKGYMDINDFGDDPIGHIISYLEK 196
QY 206 SRVYHONGBRNPHVYQOLLEGGEETLRLGLERPOSYLYLVKQCAKV---SSI-ND 261
DB 197 SRVYKOHVERNHFAYQOLLRGSEDKQLHELHERPAYNTHOGAGLMTVHSLDSD 256
QY 262 KSDMKVMKALSIYDTEDEVEDLISVAVLHGHIFPAEDSDMAO-----VTTENDL 316
DB 257 EDSHQAVTAMRVIGSPREESVHRILAAIHLGNIEVEVEEGLOEGIAVADEALV 316
QY 317 KVLTLRLLGVEGTLREALTHRKIIAKGEELLSP-LNLEQAAVARDALAAVSRFTWLV 375
DB 317 DHVAELTATPRDLVRLSLATFVAGSGRELIEKHTAAEASVARDACAAVQRLFEWV 376
QY 376 RKINSLSAKDAESWSRSTVILGLDITGEFEVPOHNSPOECICINCKEKLQDLIELT 435
DB 377 NINSVMEPRGRDPRDQKDTYIGVLDYGEFEVFNSEEDCINVCHEKLOQLPIQIL 436
QY 436 KSEOEYEAEGLIAMEVQYFNKKIICDLVEEKFGLISILDEECLRPGEATDPLLEKLE 495
DB 437 KQOEYEEYREGITGQSVFENFATVLDVERPHRILAVLDACSGACTIDRITQLD 496
QY 496 DTVKPHPHLTKLADOKTRKSLDNG-EPRLIHYGEVYTSYTGFLDKNNDLFRNLKET 554
DB 497 THRHHLHTSKQLC--PTDKTMEGRDFRKHYAGDVYTSVEGEIDKNRFLFODFFRL 554
QY 555 MCSSNNPIMAOCF-----DKSELSDKKRPETVATOFKMSLLQVLEILSKPEAVTRCKP 609
DB 555 LVNSTDPTLRAMWPRGQDITEVT--KRPLTAGTLEFKSMALVNLASKPEFYKCIKP 612
QY 610 NDAKQPRPDEVILHOKYKYLGLMENLVRRAGFAVRRKRYEAFIQRYSLCPETWP--MMA 668
DB 613 NEDKVAAGKIDENHCHHOVAYLGLLENVVRAGFASRQPSRELLRYKMTCEYTPMHL 672
QY 669 GRPOGVAVLVRLHLYKPEEYKMGRTKTEIRPKTLFATEDSLVRRQSLATKIOAMRG 728
DB 673 GSDKAASALBQHGLO-GDVAFGHSKLFINSPTLVLEDS-----713
QY 713 FHMROKFLVYKSAICISQSWMGTLGRRAKARKMAAOTIRRLIGFILRHSPRENAF 788
DB 714 ---RARLPI--IVLLLOKAWRGTLARMR-CRLRLAITIMRMR-----RIKVA-----758
QY 759 FLUDHRAFLNLRQLPRNVLDTSWPPPALRASELLRELCKMKNWYKCSISEPW 848
DB 759 HLAELQRRF--QAAQAPLYGRDLWPLRPVAVLOPFOPTCHALFCRWRAROLVKNIPSD 816
QY 849 KOOLQOKAVASEIFGKKDNY--POSVRLFISIRLIGEEISPRVLOSL-----GSEP 899
DB 817 MPOIAKAVAAAGALQGLKQDWCGRMAARDYLSATDNPTASSLFAQLTKLRDKDEGA 876
QY 900 IOYAVPVVYKRGYKPRRQLLTPSAVVVE--DAKYQRIDYANLTGISVSLSDS 956
DB 877 VLFSSHVRKVNRR-FHKIRNALLLTDQHLKYLDPDROGRVRAVPLEAVTGLSTVSGDQ 935
QY 957 LFLVLIHQREDN 967
DB 936 LVLVLIHQDOD 946

```

RESULT 9
 AAU97544
 ID AAU97544 standard; Protein; 1098 AA.

AAU97544;

13-AUG-2002 (first entry)

Human Myosin-1F protein MYO1F.

Human; phosphoinositide-binding protein; phosphatidyl acid; MYO1F; phosphatidic acid; signal transduction; housekeeping; myosin-1F.

Homo sapiens.

WO200218946-A2.

07-MAR-2002.

23-AUG-2001; 2001WO-GB03791.

23-AUG-2000; 2000GB-0020833.

15-DEC-2000; 2000GB-0030637.

(BABR-) BABRAM INST.
 (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.
 (LIMZ/) LIM Z.

Stephens L, Hawkins PT, Holmes AB, Manfava M, Kistakis N;
 Thuring WJF;
 MPI: 2002-434908/46.
 N-PSDB: ABK52205.

Novel probe for use in assay method for detecting, measuring, identifying and/or isolating PA- and/or PIPn-binding protein in a test sample, has immobilised phosphatidyl acid attached to solid support -
 Disclosure; Page 134; 163pp; English.

The present invention relates to a new probe comprising or consisting of an immobilised phosphatidyl acid derivative attached on to a solid support, or a phosphatidic acid (PA) functionalised solid support. The probe of the invention is useful in an assay method for identifying and/or isolating a protein that binds to the probe. The invention is also useful for detecting, measuring, identifying and/or isolating more than one type of phosphatidic acid and/or phosphoinositide-binding protein (PIPn) from a test sample e.g. a tissue of tissue culture extract. The assay is also useful for detecting, measuring, identifying and/or isolating phosphatidic acid and/or phosphoinositide-binding proteins in a test sample, to detect and/or measure the ability of an agent, applied to phosphatidic acid and/or phosphoinositide-binding proteins in a test sample, to detect and/or measure the ability of an agent, applied to phosphatidic acid and/or phosphoinositide-binding proteins containing detect and/or measure the ability of an agent, applied to the probe, to agonise or antagonise protein-probe binding, and to identify an agonist or antagonist of phosphatidic acid/ phosphoinositide-binding protein-phosphatidic acid/ phosphoinositide interaction, and in a single step high throughput screen of candidate agonist and/or antagonist. The invention is also useful for identifying important proteins for signal transduction, housekeeping and diagnosis. The probe is useful as an important research tool in fundamental research for diagnostics and drug discovery. The present amino acid sequence represents the human myosin-1F protein MYO1F of the invention.

Sequence 1098 AA.

Query Match 31.8% Score 1714.5; DB 23; Length 1098;
 Best Local Similarity 38.1%; Pred. No. 1.1e-136;
 Matches 400; Conservative 152; Mismatches 304; Indels 193; Gaps 23;

```

QY 28 GVDDFVLENTSEAFIENLRRENTLITYIGVLYSVNPRDQIYSROHMERRG 87
DB 18 GKDFEVLLOPIT-EDAIANLRRRREMDYITFYIGVLYSVNPFKQMPYFTDREIDLXOG 76

```

QY	88	VSFEVPHLEFVADVTRALTRERPROAVI	SEESAGKTEARKRLLOFEALFCAPER	147	
Db	77	AAQYENPHIYALTDNMTRNMLDCEMOCY	IISEESAGKTVAAKIMGYISXKSGGEEK	136	
QY	148	GGAVRDLLOSNPVLEAFGNAKTLRNDSS	RFGKTYMDQEDFGAPVGGHILSYLLEKSR	207	
Db	137	VQHAKTIDILISNPILLEAFGNAKTVRNN	SSRFGKTYFPEIOSRGEPOGKISNPLEKSR	196	
QY	208	VYHONHEERNHYVYOLLEEGEETLRILG	IERNPDSLYTVGQCAKVSISNKPDMKV	267	
Db	197	VVMONERENRHHIYOLLEGASQSOROULG	-MPDYITYLNQSDTYQVDSITDRSDPGE	255	
QY	268	MKRALSYIDFTEDEVEDLLSTVASVYLH	GINHFADEDSN-AQYTENOLKYLTRLLGVE	326	
Db	256	TLSAMQYIGIPPSIQOVLQVLAIGILH	GINISFSC-EDGNARVESVDLLAFPAVLLGID	313	
QY	327	GTLIREALTHKIIA----	KGELLSPNLQOAAVPDALAKAVYSRFETLVAKINRSL	382	
Db	314	SGRLQEKILTSKMDSRNGGKRESINVT	LVNQQAAITRALKAGIYARLDFEVLAINRAM	373	
QY	383	ASKDAESBSWSTVILGLIDYGFVEVQHS	FEQFCINYNCEKILQOULEIELTLKSEDEY	442	
Db	374	-QKQOEYS-----	IGVILYGEIRIQKNGFEQFCINFNVEKILQOULEIELTLKAEOREY	426	
QY	443	EABEIANEPVOYFNKKIICDLVEKKF-	-GIISILDECL--RPGATDILFLEKLEBDT	497	
Db	427	VQEBIRMPIDYFNKKVYCDLLENKLSPE	GMASVLDVCAFMHATGGADOTLQAKLOAA	486	
QY	498	VKPRPHETLHLDOKRTRKSLDRGEFBL	HYAGEVYTSVTEFLDKNNDLERNKETGCS	557	
Db	487	VGTEHEF-----	NSWSAG-FVIHNYAGKVSXDVSFCERNRDLVETSDILEMOT	534	
QY	558	SMNPIMACOPKSELSDKK-RPEIYADQ	PKFSLDQVLEILSKKEAYIRCIKPNDAKPG	616	
Db	535	SEQAFELRMLFPEKLDGDKRGPSTAGS	KIKKQANDLVATLRCSTPHYRICIKPNKTRAR	594	
QY	617	RFDEVLIRHQVYLGIMENLIVRAGAEV	ARRKYEAFLORYKSLCPEWPMMAGRPODVA	676	
Db	595	DMEENRKHQVEYIGILKENIVRRAGAE	VYRRQPAKFLQRYAIIIPBETPRMRGDERQVQ	654	
QY	677	VLVNHLGKPEEYKMGRTKIFIRPKTIL	FATEDSLFVNRQSATATQAAKMGFINRQKFL	736	
Db	655	HLRAVNNRPOYQWGSTKVFVKNDES	LLEEVNERKFDGFAKTQKAW-----	704	
QY	737	RVKRSALICQSWMRGTLGRKRAKKAQA	OTIRLIRGFIIRHSPRCENAFELDHVRAS	796	
Db	705	-----	RHHVAVKY-----	713	
QY	797	FLNLIRROLPRNVLDTSWPTPPRALRE	ASSELIRELCMKNNWKYCRSISPENKQOLOOKA	856	
Db	714	-----	EEMKR-----	718	
QY	857	VASEIFKKKNDYPOSVRPLFSTRPGLTE	EISPRVLOSQ-SEPIOYAVPVYKDYRKGYK	915	
Db	719	EASNILNKERRRNRISNRNFVGDYLG	LEF-RPELRQFLGKKERNDPDSYTKYDR-FK	776	
QY	916	PRPROLLTPSAVVIVEDAK-----	-VKORIDYANLNGISVSSLSDSLFLVLAHV	962	
Db	777	PIKRLDILTPKCVUYVIGREKKMKKGE	KQVCEVULKKKYDIALRBSVYSTRODDFFIL--	834	
QY	963	QREDDKQKGDVVLQSDHYIEPLTKP--	ALSADRYNNININGOSTFPA-----	1007	
Db	835	-QED-----	-AADSFLESVEFKTEFVSLCKRFE	BEATRRPLTFSDTLOFRVKKBSGW	884
QY	1008	GGPRGDIIDFTSGSELLITRAKNGHLAV	1036		
Db	885	GGGGRSVYFSRGGFDLAVLKVGGR	TLTV	913	
RESULT 10					
AAU23128					
AAU23128					
standard; Protein; 909					
AA.					

PR		12-SEP-2000	2000US-0231366
PR		14-SEP-2000	2000US-0232297
PR		14-SEP-2000	2000US-0232298
PR		14-SEP-2000	2000US-0232399
PR		14-SEP-2000	2000US-0232400
PR		14-SEP-2000	2000US-0232401
PR		14-SEP-2000	2000US-0232402
PR		14-SEP-2000	2000US-0233063
PR		14-SEP-2000	2000US-0233064
PR		14-SEP-2000	2000US-0233065
PR		14-SEP-2000	2000US-0233422
PR		21-SEP-2000	2000US-0234423
PR		21-SEP-2000	2000US-0234474
PR		21-SEP-2000	2000US-0234977
PR		25-SEP-2000	2000US-0234998
PR		25-SEP-2000	2000US-0234998
PR		26-SEP-2000	2000US-0235484
PR		27-SEP-2000	2000US-0235834
PR		27-SEP-2000	2000US-0235834
PR		29-SEP-2000	2000US-0236127
PR		29-SEP-2000	2000US-0236127
PR		29-SEP-2000	2000US-0236167
PR		29-SEP-2000	2000US-0236369
PR		29-SEP-2000	2000US-0236370
PR		02-OCT-2000	2000US-0236602
PR		02-OCT-2000	2000US-0237037
PR		02-OCT-2000	2000US-0237038
PR		02-OCT-2000	2000US-0237040
PR		02-OCT-2000	2000US-0237040
PR		13-OCT-2000	2000US-0239337
PR		13-OCT-2000	2000US-0239337
PR		20-OCT-2000	2000US-0240960
PR		20-OCT-2000	2000US-0241121
PR		20-OCT-2000	2000US-0241181
PR		20-OCT-2000	2000US-0241186
PR		20-OCT-2000	2000US-0241787
PR		20-OCT-2000	2000US-0241808
PR		20-OCT-2000	2000US-0241808
PR		20-OCT-2000	2000US-0241826
PR		01-NOV-2000	2000US-0244617
PR		01-NOV-2000	2000US-0244617
PR		08-NOV-2000	2000US-0246475
PR		08-NOV-2000	2000US-0246475
PR		08-NOV-2000	2000US-0246476
PR		08-NOV-2000	2000US-0246477
PR		08-NOV-2000	2000US-0246478
PR		08-NOV-2000	2000US-0246523
PR		08-NOV-2000	2000US-0246524
PR		08-NOV-2000	2000US-0246525
PR		08-NOV-2000	2000US-0246526
PR		08-NOV-2000	2000US-0246527
PR		17-NOV-2000	2000US-0249207
PR		17-NOV-2000	2000US-0249207
PR		17-NOV-2000	2000US-0249209
PR		17-NOV-2000	2000US-0249210
PR		17-NOV-2000	2000US-0249211
PR		17-NOV-2000	2000US-0249212
PR		17-NOV-2000	2000US-0249213
PR		17-NOV-2000	2000US-0249214
PR		17-NOV-2000	2000US-0249264
PR		17-NOV-2000	2000US-0249265
PR		17-NOV-2000	2000US-0249297
PR		17-NOV-2000	2000US-0249297
PR		17-NOV-2000	2000US-0249298
PR		17-NOV-2000	2000US-0249299
PR		17-NOV-2000	2000US-0249300
PR		01-DEC-2000	2000US-0250160

PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0253678.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Barash SC, Ruben SM;	
PI	WPI; 2001-465566/50.	
DR	N-PSDB; AAS40598.	
XX		
XX		
PT	Novel polypeptides and polynucleotides useful for diagnosing,	
PT	preventing, treating neural, immune system, muscular, reproductive,	
PT	pulmonary, cardiovascular, renal, proliferative disorders and cancerous	
PT	diseases -	
XX		
PS	Claim 11; SEQ ID NO 1124; 1180bp; English.	
XX		
CC	The present invention relates to the isolation of novel human enzyme	
CC	polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences	
CC	encoding them. The enzyme polypeptides of the invention may comprise the	
CC	functional classes of oxidoreductases, transferases, hydrolases, lyases,	
CC	isomerases or liases. The sequences of the invention are useful in the	
CC	diagnosis, treatment, prevention and/or prognosis of a wide range of	
CC	disorders including hyperproliferative disorders (e.g. cancer),	
CC	immunodeficiency disorders (e.g. AIDS) autoimmune disorders	
CC	(e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),	
CC	metabolic disorders (e.g. phenylketonuria), inflammatory disorders	
CC	(e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),	
CC	blood-related disorders (e.g. haemophilia), reproductive disorders	
CC	(e.g. infertility) and infectious disorders (e.g. Influenza). The	
CC	polynucleotides of the invention can also be used in gene therapy.	
CC	AAM22915-AAM23814 represent the novel human enzyme polypeptides of the	
CC	invention.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
CC		
XX		
SO	Sequence 909 AA:	
	Query Match 31.2%; Score 1685; DB 22; Length 909;	
	Best Local Similarity 43.7%; Pred. No. 2,7e-134;	
	Matches 376; Conservative 133; Mismatches 292; Indels 60; Gaps 18;	
OY	28 GVQDFVLLLENTSEAFETENRRRREMLITTYIGPVLYSNPRLDLOIYSNOMERYRG 87	
DB	: : : : :	
	18 GKRPDVLLDDOYTMEDFMRNLQLRERKRIITYTIGEVLYSNPYDELPLGEALARAQG 76	
OY	88 VSFEEVPRLTAADVTVRALRTERRDOAWMISGSGAGKTAEATRLLOFAETCPAPER 147	
DB	: : : : :	
	77 RELTERPRPHLAAVNAAKAKKAHNSRDICITYISGSGAGKTAEASHNIQIYAATNPBOR 136	
OY	148 GGA--VRDRLQSNDPLEAFGNNAKTLNRDNSSRFEGKYMDVOFEKGARVGHIILSYLLEK 205	
DB	: : : : :	
	137 AEVEVKRVKLVLLIKSTCVLEAFGNARTNRNHNSSRFEGKYMDINDFDKPIDGHISHYLEK 196	
OY	206 SRVYNHONGERNFHFVFOLLGGGEERTLRGLGLENPQSYLYXKGCQAKY---SST-ND 261	
DB	: : : : :	
	197 SRVLKQHNGEKNFFAFQLLGSEDKQHLHELNRPVVNYFTHQAGLANMTVSALSD 256	
OY	262 KSDKVMKNAKLAVSIDTFEDEVEDLLSYASVLHGNIFFADEDSMAQ-----VTTENQL 316	
DB	: : : : : :	
	257 EQSHQAVTEAMRVYIGFSPEEYESVHRILAATLIHGNIFFETBESGJCKDEGLAAVAERLV 316	

```

QY 317 KYTLRLGVEGTLREALTHRIIAKGELLSP-LNLEQAAAYARDAKAAVYSRTFTLV 375
DB 317 DHVAELTAPRDVLRSLLARTVAGGRLLEKHTAAEASAYARDAKAAVYQRLFEVW 376
QY 376 RRTNRSIASKDAESPWSRSTVGLGLDITYGEFEVFOHNSFEQPCINCYCKEKLQOLEIETL 435
DB 377 NRTNSVMERGRDPRDCKDITYIGVLDITGEFEVFNVSQPCINCYCKEKLQOLFTQIL 436
QY 436 KSEQERYEAGIAMEPVQYFNKKIICDLVEEKFKIISILDECIIRGEADDTLEKLE 495
DB 437 KOEOERYEGITWQSVSEYFNNAJTVDLERPHRILAVLDEACSSAGITDRILQTL 496
QY 496 DTVKPPHPLTHKLADQKTRKSLDNG-EPRLHYAGEVYSVYTGFLDKNNDLFPNLEK 554
DB 497 THRHHLHTTSKQLC--PDKTMEGRGRPRIRHYAGDVTYSEGIIDKRDPLDFDKRL 554
QY 555 MCSMNPIMAOCF----DKSELSDKRPETVATQFKMSLQVLVEILNSKEPAYIRCIKP 609
DB 555 LVNSTDPTLRAMMPDQODITEVT--KRPLTAGTLFKNSMVALVENLASKPEFYVICIKP 612
QY 610 NDAKOPGRDEVLIRHOVYKLGIMENLRVRAGFAYRRKYEAFLQRYKSLCEPTWP-MAA 668
DB 613 NEDKVAAGKLDEHNCRRQAVYIGLLENVRRAGFASRQPSRFLRYKTKCTEYTPNHL 672
QY 669 GRPDGVAALVRHLGKPEEYKMGRTKIFIRPKTLFATEDSLEVRROSLATKIQAAWNG 728
DB 673 GSDKAVALSLDQHQIG--GDVAFGHSKFLIRSPRLVLEQS----- 713
QY 729 FHMROKFLVYKSAICIQSMWNGTUGRRKAARKMAAQTIRRLIGFILLRSPCEPNAF 788
DB 714 ---RAFLIPi--IVLLQKAMRGTLAKMR--CRRLAITYIMRFR---RHKVRA----- 758
QY 789 FLDVRSASFLMLRQLPRNVLDTSWPPRPALREASELLRELCCKNMVWKCRSISPEW 848
DB 759 HLAELQRRF--QAAKRPPLIGRDLYWPLRPAYLQRPDCHALFCGRMARQLYKNIIPSD 816
QY 849 KOOLQOKAVASEIFKGRKNDY 869
DB 817 MPQIKAKVAMGALQGLRQDW 837

RESULT 11
AB97258
ID ABB97258 standard; Protein; 1089 AA.
AC ABB97258;
XX
XX
DT 27-JUN-2002 (first entry)
XX
DE Novel human protein SEQ ID NO: 526.
XX
XX Human; antihaemic; vulnerary; antiinflammatory; immunomodulator;
KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparinsonian; protein therapy; Est;
XX expressed sequence tag.
XX
XX Homo sapiens.
XX
XX OS
XX PN MO20022660-A2.
XX
XX PD 21-MAR-2002.
XX
XX PF 10-SEP-2001; 2001MO-US26015.
XX
XX PR 11-SEP-2000; 2000US-0659671.
XX
XX (HYSE-) HYSEQ INC.
XX
XX PA
XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX DR WPI: 2002-292408/33.
XX
XX N-PSDB; ABB97258.

```

```

XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
XX Example 2; SEQ ID NO 526; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate haemostasis or thrombolysis e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention.
XX
XX Sequence 1089 AA:
SQ
Query Match 29.0%; Score 1565; DB 23; Length 1089;
Best Local Similarity 35.7%; Pred. No. 6.36-124;
Matches 378; Conservative 147; Mismatches 283; Indels 252; Gaps 25;
QY 70 PYRLQIYSROMERYGVSEYFPHLFAVADTVYALATERDDAVMISGSGAKTE 129
DB 2 PY-----FTDREIDLQYQAAQYENPPIHYALTDMYRNMLIDENQCVISGSGAKTV 56
QY 130 ATKRLQYAECPAPRGAVNRDLQSNVLEAFNNAKTLRNDNSRFQKYMDVQDF 189
DB 57 AAKYIMGYISVSGGEVQVHVKIILQSNPLLEAFGNATVRRNNSRGRKFKELQFSR 116
QY 190 KGAPVGHILSYLLEKSRVYHONGERNFHYQL----- 224
DB 117 GGEPDGKISNLEKSRVYHONGERNFHYQL----- 224
QY 117 GGEPDGKISNLEKSRVYHONGERNFHYQL----- 224
DB 225 -----LEGEETLRLRIEENPQSYLIYKGCQAKY 256
QY 177 GHRADGEGSGLSQASTRCMMNAPSSWLEASQSOEORONIGL-MTPDYTYLNDSDTYQ 235
DB 257 SSINDSDMKVYRKALSYIDFTEDEVEDLISVAVLHGINHRADEDSN-AQVTTENQ 315
DB 236 DGTDRSPFGELTAMQVIGISIQQL-VIQLVAGILHGINISRC--EDGVAVAVESVDL 291
QY 316 LKYLTLRLLGVEGTLREALTHRIIA-----KGEELSPNLDEQAAVARDALAAVYSTRF 371
DB 292 LAFPAYLLIGIDSGRLQELTISRKMDSRWGSSESINVTLVNEQAAVYRDALAGLARLF 351
QY 372 TWIVRKINRSIASKDAESPWSRSTVGLGLDITYGEFEVFOHNSFEQPCINCYCKEKLQOLEI 431
DB 352 DFLVEAIRKAM-QRPQETYS-----IGVLDITYGEFEVFOHNSFEQPCINCYCKEKLQOLEI 404
QY 432 ELTLKSEQERYEAGIAMEPVQYFNKKIICDLVEEKFK--GIISILDECI--RGEAT 486
DB 405 ELTLKAEQERYEAGIAMEPVQYFNKKIICDLVEEKFK--GIISILDECI--RGEAT 486
QY 487 DITPEKEIDTVYKPPHPLTHKLADQKTRKSLDNGEPRLHYAGEVYSVYTGFLDKNNDL 546
DB 465 DQTLQKQLQAAVGHNEH-----NSWSAG-FVTHHYAGKVSIDVSGFERKRDV 512
QY 547 LFRNLKETMCSMNPIMAOCFQKSELSDK-RPETVATQFKMSLQVLVEILNSKEPAYIR 605
DB 513 LFSDLIELMOSSDDAFLMFLPEKLDGDKRPRSTAGSKIKQKQANDVATLMROCTPHYIR 572
QY 606 CIKPNDAKOPGRFDEVLIRHOVYKLGIMENLRVRAGFAYRRKYEAFLQRYKSLCEPTWP 665
DB 573 CIKPNETKRPRDMEENRYKHQVEYLGKENIRVRAGFAYRRKYEAFLQRYKSLCEPTWP 632
QY 666 MAAKRPQGVAVLYRHLGKPEEYKMGRTKIFIRPKTLFATEDSLEVRROSLATKIQAA 725
DB 633 RMRGDERGOVHLLRAVNMEDDOYQMSGTAKFVKNPSLLEVEVRKRRKDGFKRTIOKA 692
QY 726 WRGFMHROKFLRVKRSALCQSMWNGTUGRRKAARKMAAQTIRRLIGFILLRSPCEP 785

```


Db 693 W-----RRHVAVRKY-----702
 Qy 786 NAFPLDHYRASFLTLNLRQLPRNVLDTSWTPPALREASLRELCMKMMWKKYCSIS 845
 Db 703 -----EEMKE-----707
 Qy 846 PEMKQLOQKAVASEIFKCKKNYPQGVPLRFLSTRGTEBISPRVLQSIG-SEPIQYAV 904
 Db 708 -----EASNITLKKERRRNSINRNFVDYGLDE-RELNLQFLGKKERVDFA 755
 Qy 905 PVYKDRGKYPKRRROLLTPRSANVYVEDAK-----VKCRIDYANITGTSVS 951
 Db 756 SVTYDRR-FKPIKRDILTTPKCYVIGREKMKKGPCKPVCCLKKKDIOALRGVSL 814
 Qy 952 SLSDSLFVLHYQREDNKQKGDVYLQSDHYETLTKT---ALSDRVNMININGSTIEA- 1007
 Db 815 TRQDFFIL---QED-----ADSFLESVFKTFEVLCKRFEATRRPLTFSD 862
 Qy 1008 -----GCPGRDGIIDFTSSGSELLTTKAKNGLAV 1036
 Db 863 TLQFRVKKEGWGCGGTRSVTFESRGFDLAVLKVGGRITLV 902
 RESULT 12
 ABG20610
 ID ABG20610 standard; Protein; 759 AA.
 AC ABG20610;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #20601.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YF.
 DR WPI, 2001-639362/3.
 DR N-PSDB; AAS84797.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT biologically for genetic disorders or other traits and to assess
 PT biodiversity.
 PS Claim 20; SEQ ID No 50969; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 QX Sequence 759 AA:
 Query Match 26.2%; Score 1412.5; DB 22; Length 759;
 Best Local Similarity 38.5%; Pred. No. 3.7e-111;
 Matches 332; Conservative 122; Mismatches 261; Indels 147; Gaps 18;
 Qy 28 GVQDFVLLNFTSEAFIENLRFRRENLIYVIGVLSVNPYRDLQIYSQHMERYG 87
 Db 10 GRPFDVLLDQVYME-DFVRNLQRFKRIYIYIGVLSVNPYQDELPGFEALARTQG 68
 Qy 88 VSFYEYPRHLFVAVDTVYRALTERRDQAVMISGESGAGKTATYRLQFYAETCPAPER 147
 Db 69 RELYERPHLYAVANAVYKAMKHSRDTCTIVISGESGAGKTASKHIMQYIAAVNPSQR 128
 Qy 148 GCA--VRDRLOSNPVLEAFNAKTLRNDSSRPFKYMVDQDFGAPVGGHILSYLEK 205
 Db 129 AEYERKVDVLLKSTCYLEAFNAKTRNRRNSSRFKRYMDINDFGDPDGIGHISYLEK 188
 Qy 206 SRVYHONNGERNHFVYQLLEGESEETLRLGLEBNPSYLYLVGQCAKAVSINDKSDW 265
 Db 189 SRVYKOHVGERNFHAFYQALD-----SDQSH 215
 Qy 266 KYMKRALSYDTEDEVEDLSTVASYLHGINHRAEDBSNAQ-----VTENQKLYLT 320
 Db 216 QAVTEAMRYIGSPSESVSRHLLAIIHLGIEVEVEEGGLQREGVLAEEALYDVYA 275
 Qy 321 RLQVGEGLTLREALHRIKAKGELLSP-LNLQDAVAYADALAKAVYSPFTWLVKRN 379
 Db 276 ELTATPRDLYVLSLARTVYASGRELIEKGRHMAASARDAKAAVYQRLFENYVNRIN 335
 Qy 380 RSLASKDAESPWSRSTVYGLDIYGEFVPOHNSFQPCINVCNEKLOQLTELTLSQ 439
 Db 336 SYMEPRGRDPRRDQDTYIGVDIYGEFVFPVNSFEQPCINVCNEKLOQLTOLKREQ 395
 Qy 440 EEEYEGTAMEPVOYENKTIQDIYEKKFKGIISLDEECURGEGATLTLEKLEDPVK 499
 Db 396 EYEREGITWQSVETFNNAITVDLVERPHRGLAVLDACSSAGTITDITFLQTLDMHNR 455
 Qy 500 PAPHFLTKH-----LADQKTRKSLDRG-EPRLLYAGEVTVYVTFGLDKNNDLFRNLK 552
 Db 456 HHLHTSKQVPRPVPVPPQVADKTEWEPGRDPRKIKRYAD-----DLK 497
 Qy 553 ETMCSSNMPVIAQC-----DKSELSDKKRPETVATQPKMSLLOLVEILRSKPEYIRCI 607
 Db 498 RMLYSTDPDTRAMPDQODITEYT--KRPLIGTTLKKNMVALVENMLAKKEPFYVCI 555
 Qy 608 KPNDAKQGRFDEVILRHQVYKYLGLMELVRYRAGFAVRRKYEAFLOQYKSLCEETPMW 667
 Db 556 KPNEDKAVAGKIDENCRQVAVYGLLENAVS----- 587
 Qy 668 AGRPODVAVLVRLGLVYPRYKMGRTKIFTRPRTLFATEDSLVRRQSLATKIQAMR 727
 Db 588 ALKEQHGL-----QGVNARGHSLKFLTRSPRTLTLLRS----- 620
 Qy 728 GFHMOKFLRYKRSALICQSWMRITLGRKAAKKAQAOTRRLIRGLRFLHSRCEBNA 787
 Db 621 -----RRLIPV--IVLLQKAMRGTLAWR--CRRLRLYITIMRFR---RHKYRA--- 665
 Qy 788 FELDHYRASFLTLNLRQLPRNVLDTSWTPPALREASLRELCMKMMWKKYCSISPE 847
 Db 666 -HLAELQRRF--QAARQPRVGLGRULVWLPVPRVAVLPQPDICHALFCMRAROLVKNIPPS 722
 Qy 848 WKQOLOQKAVASEIFKCKKNY 869


```

XX 31-MAR-2000; 2000WO-US08621.
PF 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
PI
PI Shimkets RA, Leach M;
XX WPI: 2000-602362/57.
DR N-PSDB; AAC76860.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11, Page 4015-4016; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antiparasitic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; anticholic; immunosuppressive;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatologic; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
XX antihydrolytic; and antineumatic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SO Sequence 512 AA;

Query Match 24.6%; Score 1330.5; DB 21; Length 512;
Best Local Similarity 53.9%; Pred. No. 26-104;
Matches 260; Conservative 82; Mismatches 129; Indels 11; Gaps 4;

QY 27 VGVDFVLLNFTSEAFIENLRPRRENLIYTYIGPVLYSVNPRDQIYSROHMERXR 86
DB 15 IGVDMVLLNFTSEAFIENLRPRRENLIYTYIGPVLYSVNPRDQIYSROHMERXR 86
QY 87 GVSFVPPHLPFAVADTYVRALRTRERDQAVMISGSGAGKEARKRLQFYAETCPAPE 146
DB 74 NNFPELSEHIFALDEARSLRQDDKDCILITSESGKTEAKLVMSYAAACGCGCA 133
QY 147 RCGAVRDLRLQSNPVLEAFGNKTLRNDSSRFKYMVDQEDFKGAPYGGHILSYLLEKS 206
DB 134 EYNOVKEQLQSNPVLEAFGNKTLRNDSSRFKYMVDQEDFKGAPYGGHILSYLLEKS 193
QY 207 RYVHONHGRNHFVYQQLLEGEETLRRLGIERPQSTLYLVKQCKAVSSINDSKPMK 266
DB 194 RYVHONHGRNHFVYQQLLEGEETLRRLGIERPQSTLYLVKQCKAVSSINDSKPMK 252
QY 267 VWRKALSVDFTEDEVEDLISIVASVLHGNTHFAADSDN---AQVTENQKLYRL 322
DB 253 TVRNAMQIYGFMDHAESESLAVAVAVLGNTEFEPERVNGLDSESKDKNELKEICEL 312
QY 323 LGVESTTLREALTHKRIAKGEBELISPLNEQAAVARDALAKAVYSRTFTVLVRKINSL 382
DB 313 TGIQDQSVLERAFSEFTVEAKQEKVSTLTINVAQAYARADALAKNLYSRLFSMLVNRINESI 372

```

```

QY 383 ASKDAESPWRSTVYGLLDIYGFVFOHNSFEQFCINCNKELQOLFELTKSEOEY 442
DB 373 KQTKVRKK-----VGVGLDIYGFELFEDNSFEQFCINCNKELQOLFELTKSEOEY 427
QY 443 EAGTAMEVQYFNNKIITDLYEKFRTIISLDEBCLRPGATDLPLEKLEDTVKPPH 502
DB 428 IREDIEWTHIDYFNNAITDILNNTNGITAMLDECLRPQVTDTEFLKINQVCATRHQ 487
QY 503 HF 504
DB 488 HF 489

RESULT 15
ABB62828
ID ABB62828 standard; Protein; 2129 AA.
XX
AC ABB62828;
XX
DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 15276.
XX
XX Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
XX
XX N-PSDB; ABL06931.
XX
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure: SEQ ID NO 15276; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins
XX (ABB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 2129 AA;

Query Match 24.6%; Score 1328; DB 22; Length 2129;
Best Local Similarity 34.7%; Pred. No. 36-103;
Matches 329; Conservative 165; Mismatches 311; Indels 122; Gaps 22;

QY 29 VQDFVLLNFTSEAFIENLRPRRENLIYTYIGPVLYSVNPRDQIYSROHMERXG 88
DB 74 VEDMTITGDL-QETITLRNQNRKAKOLITTYGSMVAVINPQILPITYNRRIOLVRNK 132
QY 89 SFYEVPPHLPFAVADTYVRALRTRERDQAVMISGSGAGKEATKRLQFYAETCPAPERG 148

```

Db 133 SIAELPHIFAIISDIAEQRLQRLKENQCVISGSGAGKTESTKLLIQLA---AISGKH 189
QY 149 GAVRDRLQSNPVELEAFGNAKTLRNDNSRFGKYMDOFDEKAPVGGHLSYLLKSRV 208
Db 190 SWIEQITIEANPIMEAFGNAKTVRNDNSRFGKXIEIRFTQGAIQGARIQQLLEKSRI 249
QY 209 VHONGBERNHFVQLEEGEETLRIGL-ERNPQSTLYLVKGOCARVSSINDKSDKV 267
Db 250 VFOSHDERNNHIFYQMLAGLSTAERERKLQEOSPOYHYLAQGCFTLPGRGADKDFAD 309
QY 268 MRKALSYDFTDEDEEDLSIVASVYLHGNIPPADEDSN---AQVTENOLKYLTRLG 324
Db 310 IRAAKVLSFKPEEWSLTLALATLHGNLRFTRTEVANLATAEIDTPTMLQRYAQLG 369
QY 325 VEGTTLREALTHRKIIANGELLSPNLQOAYADALAKAVYSRTFTWLVRKINRSIAS 384
Db 370 IPISALNALQORTIFVGEHYTSLSKKAIEGRDAFVKSLYDQIFVRIYRIINETI-N 428
QY 385 KDAESPMSRSTVGLDITIGFEVYQHNSFEOFCINCEKLOQLFELTLKSDQEEYA 444
Db 429 KOVDOP---MNSIGVLDIFGFENPDNSFQOLCINANENIQQFVGHIFKMEODEXON 484
QY 445 BGIAMEPYOYFNKLTICDVEEKFIIISILDECLRPEATDLEFLEKLEDTVAPHPHF 504
Db 485 EHINMOHIEFDNOQIIDLIGKPMNIMSLIDESKFP-KGTDQTLLEKLH----- 534
QY 505 LTH---KLADOKTRKSLDGEFRLHYAGEVTVSYTGFLDKNDLLFRNKETMCSM 559
Db 535 VOHGNSIYVKGTQTSI---FGIRHYAGVVMYNPGLFLEKNDSFSGLRTRYORST 590
QY 560 NPIMAOCPKSLSD---KKRPETVATOFKMSLOLVELIRSKEPAYIRICIRPNDAKOPGR 617
Db 591 NKLYVDIPHEMPMDTAKOP-TLCVKFRNSLMDLMTLSQAHPTFIRCINPEYKEPKN 649
QY 618 FDEVLIRHOVKYTLGIMENLRVRAGFAVRRKYEAFLQRYKSLCPETWPMAGPODGYAV 677
Db 650 FDKELCVRLRYSGMETARIRRAGYPIRHAYRAVERYRLVPPVGPL---EQCDCRK 705
QY 678 LVYRLGY---KPEYKMGRTKIFTRFEKTLFATED--SLEVRQSLATK---IQAMR 727
Db 706 LARQICEVALPADSDROYGKTKFLR-----DEDDASLELQRSOLMKISVTIYORIGIR 758
QY 728 GFHMROKFLVRKSAICIQSMWRGTLGRK-----AAKRWAAQTIR-- 770
Db 759 RYLFRRYMKRYREAITTYQRYWRGRLOQRKQYVMROGFHRLGACIAAQLTTKFTMVRCR 818
QY 771 -----LINGFTLRHSPPCENAFPLDHVRASFLNLRQOLPRNVLDTSWPTPPALREA 824
Db 819 TIKLQALSRGYL-----VRKDFOKKLLERRKQONQKKEELLKAKMKEA 862
QY 825 SELLRELCKMKNVMWYKCRSISPWKQOLOQKAVASEIFKQKKNYPOSVPRLFISTRLGT 884
Db 863 EELLR-----LQQLKEOK-----EREORQOEKRLQEOERLKA 895
QY 885 EELISPRVLOSLGSEPIQYAVPVYKYDRKGYKPRROLTLTPSAVVIY 931
Db 896 EAAARNALMAAVQOKRRTKPVKQEAEPKAPTLOARNSLPPPTLTIY 942

Search completed: July 7, 2003, 14:21:20
Job time : 105.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:19:35 ; Search time 35.4566 Seconds
(without alignments)
866.342 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398

Sequence: 1 MRYRASALGSDGVRYTWEA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1191	22.1	1939	4	US-09-310-187A-1
2	1146	21.2	1120	4	US-09-147-404-1
3	1105	20.5	2548	4	US-09-172-422-1
4	1018	18.9	1886	4	US-08-938-105-3
5	160	3.0	1809	3	US-09-012-515A-12
6	160	3.0	1809	4	US-08-360-144A-12
7	160	3.0	1809	4	US-09-012-504A-12
8	160	3.0	2549	4	US-08-471-112A-3
9	160	3.0	2549	5	PCT-US95-06722-12
10	152.5	2.8	103	4	US-08-905-223-395
11	130	2.4	1529	4	US-09-134-001C-3945
12	126	2.3	567	4	US-09-134-001C-3762
13	124	2.3	1093	5	PCT-US93-03077-1
14	123.5	2.3	1333	3	US-09-356-952-2
15	122.5	2.3	1319	2	US-08-290-731C-2
16	122.5	2.3	1336	2	US-08-290-731C-6
17	122	2.3	652	4	US-09-438-833-5
18	122	2.3	813	4	US-09-438-833-12
19	122	2.3	826	1	US-08-785-241-6
20	122	2.3	826	2	US-08-480-473B-2
21	122	2.3	826	3	US-08-915-213-2
22	122	2.3	826	4	US-09-148-547-2
23	122	2.3	826	4	US-09-235-217-2
24	122	2.3	826	4	US-09-380-662-23
25	122	2.3	826	4	US-09-438-833-1
26	122	2.3	826	5	PCT-US96-10251-2
27	122	2.3	1657	1	US-08-287-959-1

28	120.5	2.2	874	3	US-08-804-439A-15	Sequence 15, Appl
29	120.5	2.2	874	3	US-08-720-229-15	Sequence 15, Appl
30	120	2.2	2154	4	US-08-841-349-4	Sequence 4, Appl
31	117.5	2.2	1151	4	US-09-134-001C-3242	Sequence 3242, Appl
32	115	2.1	1200	3	US-08-840-006-5	Sequence 5, Appl
33	113	2.1	1151	3	US-08-840-006-6	Sequence 6, Appl
34	112	2.1	810	1	US-08-785-241-7	Sequence 7, Appl
35	111.5	2.1	1181	1	US-08-488-940-2	Sequence 8, Appl
36	111	2.1	1066	4	US-09-541-782-8	Sequence 8, Appl
37	111	2.1	1194	2	US-08-723-820-8	Sequence 8, Appl
38	111	2.1	1194	2	US-08-488-940-1	Sequence 8, Appl
39	110	2.0	787	1	US-08-574-763-2	Sequence 2, Appl
40	110	2.0	1786	4	US-08-973-462-8	Sequence 2, Appl
41	108	2.0	805	2	US-08-480-473B-4	Sequence 8, Appl
42	108	2.0	805	3	US-08-915-213-4	Sequence 4, Appl
43	108	2.0	805	4	US-09-235-217-4	Sequence 4, Appl
44	108	2.0	805	5	PCT-US96-10251-4	Sequence 4, Appl
45	108	2.0	1280	2	US-08-583-276-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-310-187A-1

Sequence 1, Application US/09310187A

Patent No. 6358751

GENERAL INFORMATION:
APPLICANT: Benichou, Gilles

TITLE OF INVENTION: Fedoseyeva, Eugenia
TITLE OF INVENTION: Involvement of Autoantigens in Cardiac

FILE REFERENCE: UCSF-090
CURRENT APPLICATION NUMBER: US/09/310,187A

CURRENT FILING DATE: 1999-05-12
NUMBER OF SEQ ID NOS: 3

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1

LENGTH: 1939
TYPE: PRT
ORGANISM: Homo sapiens

US-09-310-187A-1

Query Match 22.1% Score 1191; DB 4; Length 1939;
Best Local Similarity 35.6% Pred. No. 5.3e-104;
Matches 281; Conservative 142; Mismatches 303; Indels 64; Gaps 18;

QY	4	RASALGSDGVRYTME-----SALTARDRY-----GVQDFVLENTSEAFIENTLR	49
DB	48	KATILSREGKVALETENGKTVYKEDQVLOQNPFPKIDMML-TFLHPPAVLPRNK	106
QY	50	RRPRENLITYGPIVLSVNPYRDQIYSROMEMRYEVPPHLEPAVADTVYRALR	109
DB	107	ERKAMNITTYGGLCVVYNPYKPLPYNAEVAAYRKRKSEAPPHIFSIDNAYYML	166
QY	110	TERRDAVMSIGESGAGKTEATKRLQFYAETCPAPEG-----GAVRDLLOSNP	160
DB	167	TRENQSLITGESGAGKTVNKRKYQYFASIAAGDGKKDNANKKGLIEDIDIANP	226
QY	161	VLEAGNNAKTLLANDSSNRGKYMDVDFKCAPVGHILSYLLSKSRVYVHONHGRNRY	220
DB	227	ALEAGNNAKTVANDSSNRGKFIIRHFGATKLSADIEFYLLSKSRVIFOLKERNYHI	266
QY	221	FYQLEGGEEETLRLGLERNPOSILYVKQCKAVSSINDKSMKRYKRALSYIDTETD	280
DB	287	FYQILSNKKPELMDLVLVTNNPYDAFYSGQE-VVASIDSSIELMAVDSAFVGLPTSE	345
QY	281	EVEDLLSYAVSLHLGNIFAADE-DSNAQVTTEQNKYTLRLGLVGEITLREALTFRKI	339
DB	346	ERAGYVKLLTGAIIMHGNMKFKQKOREQAEEDGEDADKSAIYLMGNSADLLGLCHPRV	405
QY	340	IAKGELLSPINTREQAAVARADAKAYYSKFTVLVKRKINSLSAKDAESPMSRTTVIG	399


```

Db      167 TDRNOSILITGESGAGKVTWTKRVIQYFAVIAIAGDRSKDQSPCKGLEDOIIQANPA 226
QY      162 LEAFNAKTLRNDNSSRFKGYMDVDFKAPGVGHILSYLLEKSRVYHONHGERNFHYE 221
        227 LEAFNAKTVNRDNSSRFKGTIXIHGATXKLASADIEYLLLEKSRVITQLAERDYHIF 286
QY      222 YOLLESGEETLRRLGLEERNPOSYLYLVKGCQAKYSSINDKSDMKYMRKALSVIDTEDE 281
        287 YOLLSKKRPEDLMDLITNNPYDYAFISQGE--TVVASIDAEELMATAIDNAFVLTGPTSE 345
QY      282 VEDLSIVASVLIHGINHRADE--DSNAQVTENOLKYLITRLGVBGTILRRLHRTKI 340
        346 KSMYVLTGAIMHFMGMMKRLKQREQAPDGTTEEDKSAVYMLGINSADLLGLCHPYXK 405
QY      341 AKGEELSPNLNEQAAVARDALAKAVYSRTFTWLVKIRNSLASKDAESPWSRTVVLGL 400
        406 VGEYEVYTKGQNVQVYATGALAKAVYERMFMMYTRINATLETOP-----XOYFICV 459
QY      401 LDYGFVFOHNSFEQFCINYCNEKIQOLFIELYLSQEQEYEAEGIANEPVOYFNKKIT 460
        460 LDYAGFEIFDFNSFEQFCINFNEKIQCFNNHMFLEQEEYKKEGIEWTFIDXGMDLQA 519
QY      461 C-DLYBEKFGIISIDECELRGGEATDTLFLKLEDTYKPRPHLTKHLDQKTR--KS 517
        520 CIDLI-EKPMGIMSLIEECMPF-KATDTEFKAKLFD-----NHLGKSANFOKPRNITG 571
QY      518 LDGGEFRLHYAGEVYTVTGFLDKNNDLFLRNLKETMCSMNPIAOCF----- 567
        572 KPEAFSLIHYAKIYVYNNIGWLQKKDKDLXETVAGLYOKSSXLSTLFAVYACADAPI 631
QY      568 DKSLSDKRRP--ETVATOFKMSLQVLEILRSKEPAYIRCIKPDOKAGRFDEVLIRH 625
        632 EKGKGAKKGSSFOYTSALHRENLMKLTNLSSTPHNEFYRCIIPNETKSPGVMDNPLVHM 691
QY      626 OYKVLGIMENLRBRAGFAVRRKYEAFLORYKSLCEPMPMAGRPDS--VAIVLIRL 682
        692 QLRGCVLEIGIRICRKGFNRLIYDXORXYILNPA--IXEGFQXSRRKXAEKLLSSL 749
QY      683 GYKPEYKMGRTKIFIRPPTLFATEDSLVARROS-LATKIOANRGFMROKFLRV--K 739
        750 DIDHNGYKFGHTKVF--FKAGLLGLLEBRKXRLSRITRIOAGRGVILAKRMEYKLLER 807
QY      740 RSAICIQSW 748
        808 RDSLVIQW 816
        Db

```

```

RESULT 3
US-09-172-422-1
; Sequence 1, Application US/09172422A
; Patent No. 6300485
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Chlu, Choi Yling
; APPLICANT: Duhil, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IIX AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; FILE REFERENCE: 200130.442
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2548
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-172-422-1

```

```

Query Match      20.5%; Score 1105; DB 4; Length 2548;
Best Local Similarity 25.2%; Pred. No. 1.4e-95;
Matches 320; Conservative 205; Mismatches 370; Indels 376; Gaps 32;

QY      40 SEAAFIENLRFRRENLIYVYIGVPLVSVNDYRDLQIYSROMHRYGVSEYEVPHLFA 99
        158 NEKTLLENLRDRFKHEKITYTVGSITLIYNPFKFLPYNPXYKVDHNLQKREPHIYA 217
Db      100 VADIVYRALRTERHDQAVMIGESGAGCTEATKLLDQFATCAFERGGA--VRDLLO 157
        218 VADAVYHMLQRRKNQOCIVIGESGSGKTSQSTNLH--HLTALSKGFASGVEQITLG 274
QY      158 SNPVLEAFNAKTLRNDNSSRFKGYMDVDFKAPGVGHILSYLLEKSRVYHONHGERN 217
        275 AGPVLEAFNAKTLRNDNSSRFKGFIOVNVYETVIGAVYEKILLESRLVYGEHNERN 334
QY      218 FHVYQVLESGEETLRRLGLEERNPOSYLYV-----YKGCQAKYSSI 259
        335 YHVFYLLAGASEDEBSAFHL-KOPEEYHYLNQITKPKRLQSWDYCYDSEPDGFTVEGE 393
QY      260 NDKSDMKYMRKALSVIDTEDEVEDLSIVASVLIHGINHRADE--EDSNAQVTENOLK 317
        334 DLRHDEKRLQAMEVGFLLPRTROTLESLSAILHLGICVKKYTRDSDIDICNPEVLP 453
QY      318 YLTRLGVEGTTLREALTHRKIIAKGEELSPNLNEQAAVARDALAKAVYSRTFTWLVYRK 377
        454 IYSELLEVKEMLEFALVTRKTVYGEKLLIPYLAELVYVYRNMASSIVSALDWIYFR 513
QY      378 INRSILA-SKDAESSWSMTYVLGLDIYGFVFOHNSFEQFCINYCNEKIQOLFIELYTLK 436
        514 INHALLNSKDEHNTP--KTLISGVLDIFGFEDYENNSFEQFCINFANERLQHYFNQIHFK 571
QY      437 SEOEYEAEGIAMEVVOYFNKKIIODVEEKFIIISIDECELRGGEATDTLFLKLEDT 496
        572 LEQETRYEGISMHNIDYIDTCCINLSKRPGLHLHLDSESNP--QATNQTLIDKFK- 629
QY      497 TVKPRPHLTKHLDQKTR--SLDRGEFRLHYAGEVYTVTGFLDKNND-----LL 547
        630 -----QHNDNSYIEPRVAMEPFIKHAGVYKVKVOKFRREKNDHMPDVAL 679
QY      548 FRNLKETMCSM----- 559
        680 LRSSKNAFISGMIGIDPVAVERMAILNFAFRVAVAFREAGKRNIRKTHGDDTAPCALTK 739
QY      560 -----NP-----YMAQC----- 568
        740 SMDSEFLQHPVHORSELTIOCKEKEYSITRKNPRTPLSDLOGMALNEKNQHDTPDIA 799
QY      569 -----KSELSD----- 574
        800 WNGRTGIRQSRLSSGTSILDDKGIFFANSTSSKLLERAHGILTRKKNKSRPALPKHLEEV 859
QY      575 -----KRPETVATOFKMSLQVLEILRSKEPAYIRCIKPN 610
        860 NSLKHFLRLQDRITRSLHLHKKKRPSSIAFOFASLSKMETLQOAEYEFVCKIRSN 919
QY      611 DAKQPRGEVLIHROYKVLGIMENLRBRAGFAVRRKYEAFLORYKSLCEPMPMAGR 670
        920 AEKLPRLRSVDYVLRQLRYTGMLETVOIRSGISSKTSFODFVSHFVLLPRNITPSKEN 979
QY      671 PDGVAVVLRHLGYKPEYKMGRTKIFI-----REPRTLFA 706
        980 IOD-----FFRKININPDYGVGKTMVFLKEGDERHODLLHQEVLRLITLLQRMFRLLC 1035
QY      707 TEDSLEVRQSL-----ATKIOANRGFMROKFL 736
        1036 RQHFLLRQASVITIQFRWNYLNOKVDAVOKDAFVMASSAALLQASRAHLERORYL 1095
QY      737 RVKRSALICIQSWMGITGLRKA-----KRWAA-----QITRLI-----RGITLR 778
        1096 ELRAAIVIQKMDYVRRRMAALICIQARKKAVRESKRYQEDQKKIITLLQSCIRGRAR 1155
QY      779 HSPRC-----PENAFFLDHVRSFLLNLRDLPRLVLDTSWPTPPALREASELL 828

```

DB 1156 QPFAKLEQRLAETPEVg--LVNTKIGSLTIOQSDSEWEDCSFDRNIRAKIECKSVI 1213
QY 829 REICMKNMVKYCRSISPEMKOOLQAKAVASEIFKGDNDYQSVPRLEISTRLGTEIS 888
DB 1214 ESNRISRESSVDCLKESPNKQOERAOSSGVLD--QED-----VLVRER----- 1255
QY 889 PVLQSLGSEPIQYAVPVKRYKRRKPPRDLPLPSAVIVEDAKYKORDYANLGI 948
DB 1256 PSLLEDLHOKKVGRA-----KRESRRMLEQATFSLLELKNR-----SLGSI 1298
QY 949 SVSISDSLFVLVHOREDNKQGDVY---LQSDH-----VIELTKTALSADRYN 995
DB 1299 SPS-----EDRMSTELVPELQSPRGTPOSSSGSLLELSTYEQSKMLE 1345
QY 996 NININGSTIF 1006
DB 1346 SVISDEGDLQF 1356

RESULT 4
US-08-938-105-3
Sequence 3, Application US/08938105
Patent No. 6353151
GENERAL INFORMATION:
APPLICANT: Ielinvand, Leslie A.
APPLICANT: Vikstrom, Karen L.
TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: CO
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,105
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crook, Mannel M.
REGISTRATION NUMBER: 31,071
REFERENCE/DOCKET NUMBER: 3595-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1886 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-938-105-3

Query Match 18.9%; Score 1018; DB 4; Length 1886;
Best local Similarity 32.1%; Pred. No. 1,8e-87;
Matches 253; Conservative 136; Mismatches 286; Indels 114; Gaps 17;

QY 4 BASALGSDGVRYTME-----SALTARDRY-----GYOQFVLENTSAAIENR 49
DB 47 KATVSRREGKVTALFENGKTYTVKEDQVMOONPKFKIEDMANI-TFLHPAVALYNK 105
QY 50 RREFRNLITYIGPVLSVNPYRDQIYSRQMERIRGVSEFVPPHLFAVADTVYRALR 109
DB 106 ERYAAMMITYIGLPCVTVNPKWLPYVAEYVAAYRGKKREAPPHIFISIDNAYQYML 165
QY 110 TERRDQAVMISGSGAKTEATKRLLOQFYAETCPAPERG-----GAVDRLLQSNP 160

DB 166 TDRENOSTILTIGESGACKTWTAKRVIOYFPASIALIGRSKNDPNANKGLLEQIIQANP 225
QY 161 VLEAFNAKTLRNDNSSRFGKYMDVQDFGAPYGHIIISYLEKSVYHQNGERNFV 220
DB 226 ALFAFNNAKTVRNNDSSRFGKFIIRHFGANGKLASADIEYTLLEKSHVIFQLAERNYHI 285
QY 221 FYOLLEGESEETLRLGIERNPQSYLYLVYGQCAKVSINDKSPMKYRKALSYIDPTE 280
DB 286 FYQLTSNKKPELMLLTNNPNFYAVVSGQE--SVASIDSEELLATDSAPFVLTAE 344
QY 281 EVEDLLSVASVLIAGNIHFAADE--DSNAQVTTENOLKYLTRLGVEGTLRLAALHRT 339
DB 345 EKAGVYKLTGAIMHYGMKKFKOKREBOAEDDGTEDAKSAYIMGLNSALLKGLCHPV 404
QY 340 IAKGEELLSPLEQAAAYARALAKAYSRFTLVYKIRSLASKRAEBSRSTVIG 399
DB 405 KVGENEYVTKGSOVOYVYSIGALAKSVYEKKFMNVRINATLETQPRO-----YFIG 458
QY 400 LLDIYGFVFPQNSFEQFCINYCNEKLOQLFIELTLKSEQEYEAEGIAPEYQYFNKI 459
DB 459 VLDIAGEF----- 466
QY 460 ICDLVEEKFGIISILDECLRPGAADLFLERLEDPYKPHPHLTHTKLADOKTR--KS 517
DB 467 ISSLPHTKIMGIMSILEECMF--KATDMTFKATLYD-----NHLGKSNFQKPRVKG 519
QY 518 LDRGEFLHYAGVTVSYNGEFLDNKNDLFLRNLEKMGSSMNPINAOCPKSELSD-- 574
DB 520 KOEAFHSLVHYAGTVNILLGLMKNPDQWETVYGLYOKSSLKMLMTLSTYASADTGD 579
QY 575 -----KKRP-----EVATQFKMSILOVELIRSEKAPYRJCIPNDAKPGREDEVILR 624
DB 580 SGKGKGGKKKSSQYTSALHRENLMNTLRTTHHFRCILPNRKAAGVMDNPLV 639
QY 625 HQVYTLGIMENLVRPAGFAFRKYEAFLOKYKSLCEPTPMAGROD---GYAVLYRH 661
DB 640 HOLRNGYLEIRICRKGFPNRIILYGDFOKRYRLNPAAP--EGQIDSGKAELKLG 697
QY 662 LGVPEEYKMKRTKIFIFPPTLPAFEDSLEVRROSLATYIOAARGFHMRQFLRY--K 739
DB 698 LIDIDNQKRGHTVFER--AGLLGLEEMRDERLSRIITRIQANARQRLRIEKKMYER 756
QY 740 RSAICIOGW 748
DB 757 RDALLVIOW 765

RESULT 5
US-09-012-515A-12
Sequence 12, Application US/09012515A
Patent No. 6127521
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,515A

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/360,144
 FILING DATE: 20-DEC-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: APV-036.02
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1809 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-012-515A-12

Query Match 3.0%; Score 160; DB 3; Length 1809;
 Best Local Similarity 19.1%; Pred. No. 2e-05;
 Matches 193; Conservative 121; Mismatches 318; Indels 376; Gaps 46;

137 FYAETCPAPERGAVER--DRLLQS-----NPVLEAFGNKTLRN---DNSS--- 177
 382 FDAPAPLPSPKKALETVDRLTESLDFDYASRIIPIVRLDQSPELRSTAMDLSLV 441
 178 -REGYMDQVDFKGPAPVGHILSYLLEKSRVYHNM-----GERNFHYF 222
 442 FOLGKKYQIFIP-----VWNVVLVRRHNRHQRVYLICRIYKGYTLADEEDPLIY 492
 223 Q--LLEGEETLRRLGLERNFOSYLVKGCACAKVSSINDKPMVKRAKLSVIDFED 280
 493 QHRMLRSGGQDALASGPVETGPMKLN-----VSTINLQKANGAARR----- 534
 281 EVEDLSTIVASVHLGNHFAADEDSNAOVTTENOLKYLTRLGVEGTTLSREALTHRKII 340
 535 -----VSKDWLEWLRL----- 547
 341 AKGELLSPNLNQAAYARDAALAKAVYSRTFTWLVRKINRSLASKADSPSRSTYVIG- 399
 548 --SLELL-----KDSPPSLRSCWALAQ 568
 400 ----LDIYGEVGFQHNSPFOFCINYCNEKLOQLF---IELTLKSEQEEYEAGIAMEP 451
 569 AVNPMARDLF-----NAAFVCSWSELNEDQOEBLRSIELALTSQD-----IA-EV 613
 452 VOYFNKKITCDLVEKFKGIISIDE-----ECLAPGEATDLTLEKLEEDTVKP 500
 614 TOTLLN--LAEFMEHSDKGPLRLDDNGIVLLGERAAKCRAYAKA--LHYKELEFGKGP 668
 501 HPNFL-----THKLADQTRKSLDRGEFRLHYAGEVTVSVTF----- 539
 669 TPALLESLSINNKI--QOPEAAGVLEYAKNH--GELEIATWYEKIHMEDEALVAYDK 725
 540 -LDKKNLDFRNLKETMCSMNPTMA-----QCFDKSELSDKRREPETAQFKMSL 589
 726 KMDTNKD-----DELMIGRMRCLEALGEMQOLHQOCCCKEKTLYN--DETOAKMARAA 777
 590 LQVLEILRSKEPVIATRCIKPNDANQPGFDEVLLRHQ-----VKYGLMENLRVR 639
 778 AAANGLCQWDSMEETTCMIPRDTHDGAIFYRAVLALHDLDFSLAOCIDKARDLLDAELTA 837
 640 RAGFAVRRRYEAFLORYKSLCPETWPMAGRPQDGVAVLVNHLGYKPEE--YKMGRTKI 696
 838 MAGESYBRAYGAMVS--C-----HMSLELEVYQYKL----- 867
 697 FIRPKTLFATEDSELVAROSLAKIQAMRGFWRMROKFLVYKRSALC-----IOSWNGT 752
 868 -----VPERREIIRQIWMERLQGCCOQRIVEDMOKILMVRSLVSPHEDMTW---- 913
 753 LGRRKAARKMAQOTIRRLNGFILRHSPRCPENA--FFLDHVRVSFLNLRQLPRVLD 811

DB 914 -----LKTASLCGKGRALAHKTVLLGVDP---PSOLD 946
 QY 812 TSWPTPPALREASELLRELCKMNMWYKCRSISP-----EMKQOLOQKAVASEIFK 863
 DB 947 HPLPTVHPQVITYA-----YKMNW-KMSARKIDAFQHMQHFVOTMQQOQHATATEDQ 998
 QY 864 GKNDYPOSVRLTIRLGTSEISPRYLOSSEPI-----QYAVPVYKIDRKGYKPRPR 919
 DB 999 HKQELH-KIWARCEL--KLGEWOLN--LOGINSEITPKVLYQVYSAATEHDRSWYKAMHA 1052
 QY 920 QLLTPSAVV-----IYDAKVQR--IDVANLGISVSSISDSLFVLHQRENNKQCD 972
 DB 1053 WAWNFEVLYHKHQNOARDKCKKLBRASGANITNATTAATTAATTAATTSSTEGSSE 1112
 QY 973 VLIQSD-----HYLETJTKT-----ALSADRVNI 997
 DB 1113 AESTENSPTSPLOKVTEDLSKTLMLVTVPAVGFPFRSISLSGNL 1160

RESULT 6
 US-08-360-144A-12
 Sequence 12, Application US/08360144A
 Patent No. 6150137
 GENERAL INFORMATION:
 APPLICANT: Berlin, Vivian
 APPLICANT: Chiu, Maria Isabel
 APPLICANT: Cottarel, Guillaume
 APPLICANT: Damagnez, Veronique
 TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESS: POLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/360,144A
 FILING DATE: 20-DEC-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: APV-036.02
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1809 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-360-144A-12
 Query Match 3.0%; Score 160; DB 4; Length 1809;
 Best Local Similarity 19.1%; Pred. No. 2e-05;
 Matches 193; Conservative 121; Mismatches 318; Indels 376; Gaps 46;

137 FYAETCPAPERGAVER--DRLLQS-----NPVLEAFGNKTLRN---DNSS--- 177
 382 FDAPAPLPSPKKALETVDRLTESLDFDYASRIIPIVRLDQSPELRSTAMDLSLV 441
 178 -REGYMDQVDFKGPAPVGHILSYLLEKSRVYHNM-----GERNFHYF 222
 442 FOLGKKYQIFIP-----VWNVVLVRRHNRHQRVYLICRIYKGYTLADEEDPLIY 492

QY 864 GKDNYPQSVPLRFLSTRGTEESPRVLOSIGSEPI-----OYAVPVYKDRKGYPRR 919
 Db 999 HKEELH-KIMARCEL-KLGEOMLN---LOGINESTIPKVLQYYSATEHDSWYKAWMA 1052
 QY 920 QLLITPSAVV-----IVEDAKYKOR-IDYANLTGISVSLSDSLFVLHVOREDNKOKD 972
 Db 1053 WAWNFEAVLHYKHONQARDEKKLRHAGANITNATTAATTAATTTASTEGSSESE 1112
 QY 973 VVLQSD-----HYETLTFT-----ALSADRVNNT 997
 Db 1113 AESTENSPPTSPLOKKTEDSLKTLMTYVPAVOGFFRSISLSRGNNL 1160
 RESULT 8
 US-08-471-112A-3
 ; Sequence 3, Application US/08471112A
 ; Patent No. 6313264
 ; GENERAL INFORMATION:
 ; APPLICANT: Molnar-Kimber, Katherine L.
 ; APPLICANT: Fallili, Amedeo F.
 ; APPLICANT: Caggiano, Thomas J.
 ; APPLICANT: Nakanishi, Koji
 ; APPLICANT: Chen, Yanguo
 ; TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/471,112A
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/384,524
 ; FILING DATE: 13-FEB-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/312,023
 ; FILING DATE: 26-SEP-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/207,975
 ; FILING DATE: 08-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sleiman, Michael T.
 ; REGISTRATION NUMBER: 36,276
 ; REFERENCE/DOCKET NUMBER: 01142, 0058-00000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-408-4000
 ; TELEFAX: 202-408-4400
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2549 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-471-112A-3
 Query Match 3.0%; Score 160; DB 4; Length 2549;
 Best Local Similarity 19.1%; Pred. No. 3.7e-05;
 Matches 193; Conservative 121; Mismatches 318; Indels 376; Gaps 46;
 137 FYAETCPAPERGAVER--DRLLQS-----NPLVLEAFGAKTLRN---DNSS--- 177

Db 1122 FPAAPAPLPSKAALETVDRLTESLDFPDYASRIIHPITVRLDQSPELRSTAMDTLSLV 1181
 QY 178 -RFGKYMDVOEDFPGAPVGHILSYLLEKSRVHONH-----GERNFHFY 222
 Db 1182 FQLGKQYQIFP-----WVKVLVRHHIINHQRVYVLCRIYKGTTLADEEDPLIY 1232
 QY 223 Q--LLEGEDEETLRRLGLERNPDSLYLVKQCAKAVSSINDKSPMKVRKALSVIDFED 280
 Db 1233 QHRMLRSGQDGLASGPVETGPMKRLH-----VSTINLOKAMGAARR----- 1274
 QY 281 EVEDLSIVASVLIHGINHPADEDSNAQVTTENOLKYLFTLLVEGTLREALTHRKII 340
 Db 1275 -----VSKDMLWLRRL----- 1287
 QY 341 AKGEELSPPLNEQAAYARDAANAVYSRTFTWLRKINRSIAKDAESPMSRSTVVG- 399
 Db 1288 --SLELL-----KDSSPSLRSCMMLAQ 1308
 QY 400 -----LDIYGFVYVPHNSFEOFCINCNELQOLF-----ITLTKSEDEYEAGIAMEP 451
 Db 1309 AYNPARDLF-----NAAFVSCWSELNEQODELIRSLTSLALTSD-----IA-EV 1353
 QY 452 VOYFNKTIICDLVEKEFGIISILDE-----ECLYRGATDLTFLEKLEDVVKP 500
 Db 1354 TOTLLN-LAEFMHSDKGPLRLDDNGLVLLGERAKCRAYAKA---LTKLELFQKP 1408
 QY 501 HPFL-----THKLADOKTRKSLDRGEFRLHYAGEVTVSVTF----- 539
 Db 1409 TPALLESLSISNNKL--QOPAAAGVLEYANKHF-GELEIQAITYEKLHEMEDALVAYDK 1465
 QY 540 -LDKNNDLFRNLKETCSSMNPIMA-----QCFDSELSDKRPREVATOFKMSL 589
 Db 1466 KMDTNKD-----DELMIGRRCLLEAGWQLHOCCCEKWTLVN---DEQAKMAWMA 1517
 QY 590 LQVLEILRSKEPAYIRICIPNDAKQGRFDEVLIRHO-----VKYLGLENLRVR 639
 Db 1518 AAAGLQGDQMSMEYTCMIPRDTHDGAFAFVAVLALHDDLSLMOQCIDKARDLLDALTA 1577
 QY 640 RAGFAVRRKTYBAFLQRTKSLCPETPWMAAGRPDQVAVLVRLHGLKPEE--YKMGRTKI 696
 Db 1578 MAGESYSRAVGAMYS--C-----HMLSELEVIQYKL----- 1607
 QY 697 FIREPTLFATEDSLEVRROSLATKIOAMRGFMHROKFLRYKRSATC-----IOSMRRGT 752
 Db 1608 -----VPERREIIRQIWMERLOGCQRIVEDMKILVNSLVVSPHEDMTM----- 1653
 QY 753 LGRKKAARKMAOTIRRLIRGLTIRHSPRCPENA-FELDHVRASFLNLRLPRNVLD 811
 Db 1654 -----LKYASLCGKSGRLALAHKTLVLLGLVD---PSRQLD 1686
 QY 812 TSWPTPPALREASSELLRELOCMKMYKYGCRISF-----EMKQOLOOKAVASATFK 863
 Db 1687 HPLPTVHPQVITYA-----YKMKM-WKSARKIDAQOMHOFVOTMOQOAHIAEDQO 1738
 QY 864 GKDNYPQSVPLRFLSTRGTEESPRVLOSIGSEPI-----OYAVPVYKDRKGYPRR 919
 Db 1739 HKEELH-KIMARCEL-KLGEOMLN---LOGINESTIPKVLQYYSATEHDSWYKAWMA 1792
 QY 920 QLLITPSAVV-----IVEDAKYKOR-IDYANLTGISVSLSDSLFVLHVOREDNKOKD 972
 Db 1793 WAWNFEAVLHYKHONQARDEKKLRHAGANITNATTAATTAATTTASTEGSSESE 1852
 QY 973 VVLQSD-----HYETLTFT-----ALSADRVNNT 997
 Db 1853 AESTENSPPTSPLOKKTEDSLKTLMTYVPAVOGFFRSISLSRGNNL 1900
 RESULT 9
 PCT-US95-06722-12
 ; Sequence 12, Application PC/TUS9506722
 ; GENERAL INFORMATION:
 ; APPLICANT:

```

TITLE OF INVENTION: Immunosuppressant Target Proteins
NUMBER OF SEQUENCES: 25
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06722
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 27-MAY-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 20-DEC-1994
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2549 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-06722-12

```

Query Match 3.0%; Score 160; DB 5; Length 2549;

Best Local Similarity 19.1%; Pred. No. 3.7e-05;

Matches 193; Conservative 121; Mismatches 318; Indels 376; Gaps 46;

```

137 FYAETCPAPERGAVER--DRLIQS-----NPVLEAFGNKTKLRN--DNSS--- 177
1122 FDAPERPAPSRKALETVDRLTESLDFDYASRIIHPRIKLDQPELSTAMDILSLV 1181
178 -REGKMYDQDFPKGAPVCGHLLSTLEKSRVYHQN-----GERNFHYV 222
1182 FOLGKKYQIFIP-----WNVKVLVRRHRIHORYDVLICRIYKGYTLADEEDPLY 1232
223 Q--LLEGGEETLRRLGRNPOSYLIVKGCACAKVSSINDKSMKVMKALVIDETED 280
1233 QHRMLASGGDALASPVEFGPKLKH-----VSTINLCKANCAAR----- 1274
281 EWEELLISIVASVHLGNIHFAADESDSNAOVTTENDIKYITRLIGVEGTTLRREALTHRKII 340
1275 -----VSRDDMLEWLRRL----- 1287
341 AKGEELLSPLENGAAYARDALAKAYSTFTFWLVAKIKIRSLASKDAESPSMRSTVYG- 399
1288 --SLELL-----KDSSEPSLSRCWALAQ 1308
400 -----LLDIYGEFVFOHNSFEQCIIVYCNKLIQOLF--TELTKSEGEYEAEIAMEP 451
1309 AYNPMARDLF-----NAFVSCWSELNEDQDELIRSELALTSQD-----IA-EV 1353
452 VOYFNKKITCDLVEEKKGISLIDE-----ECLRGGEATDLTFLEKEDTVAP 500
1354 TQTLIN--LAEFMEHSDKGPLRLDNGIIVLGERAKARAYAKA--LHYELEFQKCP 1408
501 HPHFL-----THKLADOKTRKSLDRGEFRLLHYAGEVYVYVYVGF----- 539
1409 TPAILSLISINKL--QOPEAAAGVLEYAMKHF--GELEIQATWYKLEHEDALVAYAK 1465
540 -LDKNDLLEFRNLKETMSSMNPIMA-----OCFDKSELSDKKREPVAATQFKMSL 589
1466 KMDTKND--DPEMLGMRCLFALGEMGLHQCCCKWTLVN---DETQAKMARMAA 1517
590 LQVEILRSKEPYIYICIRPNDAKQGRDEVILIRQ-----VYLLGMEWLRAR 639
1518 AAAMGQWDSMEETCMIPRDTHDGAFFRAVALHODLFLSAQOCIDARDDLDELTA 1577
640 RAGFAAYRRKYEAFLQRYKSLCPETWPMAGRPQDGAVALVRHLGYKPEE--YKMGRTKI 696
1578 MAGESTSRAYGANVS---C-----IHLSTLEEVYQYKL----- 1607

```

```

QY 697 FIREPKTLFATEDSELYVRQSLATKIQOANRGCFHMRQKFLRKVRSACI-----IOSWMRGT 752
DB 1608 -----VPEREILRIOWMERLOGCQRIYEDMQKILMVRSLVSPHEDKRTV----- 1653
QY 753 LGRRAKARRKWAQTIIRLLIRGFLIRHSPCPENA--FLDHYVASFLINRRLPRNVLD 811
DB 1654 -----LKYASICGSGRLALAHKTLVLLGLVD---PSROLD 1686
QY 812 TSWPTPPALREASELLRELOCMNMVWKYCRSISP-----EMKQOLQOKAVASEIRK 863
DB 1687 HPLPTVHPQVYTA-----YMKNM--WKSARKIDARQHMOPVQTMQQAQAHAIATEDQ 1738
QY 864 GKRDYPOSVPRLEFTRIGTEISPRVLOSLEPI---QYAVPYVYKRYKGYPRPR 919
DB 1739 HKQELH-KIMARCL--KLGEMQULN---LOGINSETPKYLYQYSAATEHDSMYKAWHA 1792
QY 920 QLLITPSAVV-----IVEDAKYQR-IDYANLTGISVSSLSLYLVHQREDNKQKG 972
DB 1793 WAVNFEAVLVHYKHQNOARDEKKLRHASGANITNATTAATTAATTTASTSGSNSSE 1852
QY 973 VVLQSD-----HYLETLTK-----ALSADRVNNI 997
DB 1853 AESTENPTSPSPLOKKVTEDELSTLMTVTPAVQGFRRSISLSRGNNL 1900

```

RESULT 10

US-08-905-223-395

Sequence 395 Application US/08905223

Patent No. 6222029

GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste D.

APPLICANT: Duelet, Aymeric

APPLICANT: Lacroix, Bruno

TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

NUMBER OF SEQUENCES: 503

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Win95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/905,223

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Isaielsen, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-0176

TELEFAX: (619) 235-8550

INFORMATION FOR SEQ ID NO: 395:

SEQUENCE CHARACTERISTICS:

LENGTH: 103 amino acids

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN

ORIGINAL SOURCE:

ORGANISM: Homo Sapiens

TISSUE TYPE: Brain

FEATURE: sig-peptide

NAME/KEY: -72...-1

LOCATION: -72...-1

IDENTIFICATION METHOD: Von Heljne matrix

OTHER INFORMATION: score 3.6

OTHER INFORMATION: seq CTSLLQLYDASNS/EM

US-08-905-223-395

Query Match 2.8%; Score 152.5; DB 4; Length 103;
Best Local Similarity 29.2%; Pred. No. 7.1e-07;
Matches 35; Conservative 26; Mismatches 16; Indels 43; Gaps 4;

QY 615 PGARDEVILRHQVYGLMELNLRVRRAGFAVRRKRYEAFLORYKSLCPETWPMAGRPOG 674
D 2 PDQFDQAVVNLQRLYSQMLETVRLRKAGAVRRPQDFYKRYK----- 44
QY 675 VAVLVRHLGKPR-----EYKMGRTKFTFPPTLATEDSLEVRQ 716
D 45 --VLMRMLAL-PEVGRGKTSLLQLYDASNSEWOLKTKVFLR-----ESLEOKLEKRR 96

RESULT 11

US-09-134-001C-3945
; Sequence 3945; Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3945
; LENGTH: 1529
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3945

Query Match 2.4%; Score 130; DB 4; Length 1529;
Best Local Similarity 18.2%; Pred. No. 0.011;

Matches 131; Conservative 100; Mismatches 213; Indels 276; Gaps 36;

QY 25 DRVGVDFVLENETSEAF-----IENLRFRRENLI-----YVIGPVLSVNPYRDQIY 77
D 49 DACGIGFVAMNDKNSHDIIEKSLERLRDLHGGVADGIGTGAGINTETIPY---DLF 105
QY 78 SHOHMERVGVSVFYEVRPHLFAVADTVYRALRTERDDQAVMISGSGAKTEA---TKRL 134
D 106 EQLTEKVPGEQYAV--GLFPSKEKVRDSIHEMFNOYFESEGFVKVIGYRVPYDTRAI 163
QY 135 LGFVATPCAPPR-----GGAVADRLLQSNPVLEAFGNMKT----- 170
D 164 AOHVADTMPYIQOVFVDITGVKEVRLEFLARKQIEKSETOQSIDLYFTSLSHRTIYVG 223
QY 171 -LRANDSRFGKYMDOVFDEKCAPGHILSYLLEKSRVYHONHERNP-----HYVQ 223
D 224 WLRSDQIK--GLYDLQNEAVOSKIG-----LVHSRSTNTFSPSKRAHPRNM 269
QY 224 LLEGGEETLRRLGLERNPOSYLYLVKQOCARVSSINDSKDWKVRKALSYIDFTEDEVE 283
D 270 LHMNEB-----INTIKGNVMMRQRQKIVETLEDEKD 303
QY 284 DLSLTVASVYLHGNTHFADEDSNAQVTTENQKLITLLIGVEGTTLEALTHRKTIKNG 343
D 304 -----KHFYVDEGSDSIYDVALEFLSLAMEPE-----KA 335
QY 344 BELSLPLNEQAAV--ARDALAKAVYSRTFTWLVKINRSLASKDAESPMSRSTYVGLL 401
D 336 AMLLIP---EPWLVNESNDKVRSPY--EFGYSLM-----PMDGPTM----- 373
QY 402 DLYGEVFOHNSFEQFCINYN--EKLOQLFTELTLKSPQEEYEAEGIAMPEVQY----- 454
D 374 -----ISFCNGDKIGAL-----TRNGL--RGRGYTITKDN 402

QY 455 ---FNNKI-ICDLVEEK--FKGIISLDECELRPGEDATDLFLE-----KLEDTVK-- 499
D 403 FTVESSEGVIVPENNVAFKG-----QINPGKLLLVDFLQKRVENNELKTNIANE 454
QY 500 -PHPHFLTKLADQTRKSLDRGEPRLHYAGEVYVSTGFLDKNNDLFLRLK----- 552
D 455 LYEQW-----LDYKKNKNDLN-----IYQSS-----DWDQTLFRLQKQFAYTK 496
QY 553 -----ETWCSNMPIMAOQFDK--SELSDKRRPETAQOFKMSLLQVLEILRSKPA 602
D 497 EDINKYMDLVYINKKDPDIGAMGYDAPIVANLNDK--PESLFNFKQLQFQVYN----- 546
QY 603 YIRCIKPDNAKPGFDEVILRHQVYGLMELNLRVRRAGFAVRRKRYEAFLORYKSLCP 662
D 547 -----PPIDA-----YREKIVYSELISGSEGNL-----LCPD 574

RESULT 12

US-09-134-001C-3762
; Sequence 3762; Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3762
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3762

Query Match 2.3%; Score 126; DB 4; Length 567;
Best Local Similarity 18.4%; Pred. No. 0.0048;

Matches 95; Conservative 106; Mismatches 172; Indels 142; Gaps 25;

QY 204 EKSRYVHONHERNP-----HVFQLEGGEEETLRRLGLERNPO----- 243
D 60 KKKETTNIAHNEKNIILIEKDKKTIYS--SGNIKIDHRIKDNEANPSKLINKNTKGM 117
QY 244 --SYLYLVKQOCARVSSINDSKDWKVRKALSYIDFTEDEVEDLSLTVASVYLHGNTHFA 301
D 118 RYTYKNTIDDKTIYISGINNE-----IIDQKD-LMKYLSIV-GIYVLFYVYLA 164
QY 302 ADEDSNAQVTTENQKLITLLIGVEGTTLEALTHRKTIKAGEBLSPLNEQAAVARDA 361
D 165 SRSINRYIRIPINENVYATSLA-----DGYVHNV-----PESVWKE----- 202
QY 362 LAKAVYSRTFTWLVKINRSLASKDAESPMSRST-----TVGLDIYGEVFOHNSFE 415
D 203 -TRALFTYT--NDLARRQKLNNSOKIOSNRLKTLLENIPSSVL-MDKHGEIYVANNIAY 259
QY 416 QF-----CINYCNEKLOQLFTELTLKSPQEEYEAEGIAMPEV--QYFNKTIICD 462
D 260 QVFNPDQMVENKSYIGFIDISIEKLITE--SFRREKVIYEDOLEVAINNVHTKPYDVSCIP 318
QY 463 LVEEK--FKGIISLDECELRPGEDATDLFLEKLEDTVKVKKHPPHFLTKLADQTRKSLDR 520
D 319 LTKSKKNLQGGVAVVLAH-----DITWLOKLEMLRREFVANVSHELTPTT----- 362
QY 521 GEERLHYAGEVYVSTGFLD-----KNNDL--LFRN--LKETWCSNMPIMAOQFDK 569
D 363 -----SINGFAETLIEGAKKNDQSLDMLNIIILDES--NRRESLYTDLDL 406
QY 570 SELSDKRRPETAQOFKMSLLQVLEILRSKPAVIRCIKPDNAKPGREDEVILRHQVY 629

LENGTH: 1333
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-356-952-2

Query Match 2.38; Score 123.5; DB 3; Length 1333;
 Best Local Similarity 18.3%; Pred. No. 0.037;
 Matches 212; Conservative 175; Mismatches 369; Indels 405; Gaps 61;

```

OY 70 PYRDLQIYRQHMERKRGVSFYEPPLFLAVADTVYRALRTERDQAVMISGSGAGKTE 129
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 PY---EFFSEBNAPKRWGL---LVPALKKVOGQVHPTL--ESNDALQYVEE----- 50

OY 130 ATRKLQFYAETCPA--PERGAVARLRLQS--NPVLE--AFGAKTLRNDSSRFCKYADY 185
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 ---LIQLMLMCOADPRASDVEEVQKSFPHIDKMAIADAQSAIEKRRKRLSLPV 107

OY 186 QDPFKGAPVGGHILSYLLEKSRVYHQNHEFNHFVYOLLEGESEETLRRLGLEERNPOSY 245
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 E---KIHPLKEVLGYKIDHVSIVY-----IVAVLEVISADILKIVG-----NY 148

OY 246 L-----YLVKGCQAKVSSINDSKVMKALSVDFTEDEVEDLSTIVASVHLGNTHF 300
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 VNRIRHETIKODIKVAMCADK-----VLMDMFHODVEDI-----NITS 187

OY 301 AADED-----SNAQVTENQDKYL-- 319
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 LTFDEPSTSGEQTYYVDLVKAFMAEIRYIRELNILIKYRFRPFVNSNLSFSANDNIFS 247

OY 320 -----TRLLG--VEGTT--LREALTHRKIIANGELLSPLNLE--QAAVARALAKAV 366
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 RIYDIHELKVLGHIHEDIVEMTDGSPHVGSCFEDLAELARDPYESARILDRGEF 307

OY 367 YSRFTWLVLR-----KINSLS--ASKDA 387
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 HDREFLSQSLRPAALYLQSIGEGFEKAVOYVLPRLLPAYHCHLHYFELLQLEKESDQ 367

OY 388 EDSFWRSTVLDLDIY--GFEVFOHNSF-----EOPCINCYNE--KIQOLFIEITLSEQ 439
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 EKECEKQALITALLVNOGSMERIKCSLAKRRLSASACRFSSQMKMGQOLAK--KME 424

OY 440 EEEVAGIAMEPVQYFNKKIICD---LVEEKFQIISILD--BECLR-----PGEA 485
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 IOKNIDMGKIDIGCCNEFIETEGTLTVNGAKHEHIFLFGMLICKSNHGQPRLGAS 484

OY 486 T-----DLTFLEKLE---DTVKRPHPLTHKLADOKTRSLDRGEFRLLHYAGEVYS 535
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 485 NAEYRLKEKEFRKQVQINDKODTNEYKHAFFELIKDENS-----VIFS 527

OY 536 VGFGLDKNN-----DLFRNLKETMSSMNPIMACCFDSSELSDDKRPETVATQFKMSL 589
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 528 AKSABEKKNNMAALISIQYRSTLERM-----LDVYMLQEEKE-----BQMR 568

OY 590 LQLVILSKSEPAVYIRCIKPNDAKOPGFDE-----VLIRHVKTLGIMENIRVRAG 642
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 569 LPSADVYFAER-----DSEENITIEENMOPKAGIPIYAGVYIKIIEIRLYHM-- 617

OY 643 FAYRRKYAFLQRYKSLCPEPTWPMAGRPDOGVAVLVRLGY--KPEYKMGRTKIFIRFP 701
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 618 YADPFVYRTFLTYRSEC-----KPOELLSLIEREIREPERETEDR----- 660

OY 702 KTLFATED-----SLEVR-----QSLATKIQAMNGEFLW--ROKFLRVKRSACIQSW- 748
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 661 ---IAIENGDOPLSAELKRFKREKEYIQVOLAIVNCR--HMVGHNFDEFERDAYLLQRM 715

OY 749 -WRSTLGRKAARKMAAQITRILIR-----GFLHSPRCPRNAFFLD---H 792
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 716 EFIGIV--RGKAMKM--VESTIKIQKKIARONGRPHNTTFOSSP--PIVENHISRPGH 770

OY 793 VRASFLNLR-----ROL-----PRNVLDTSWP-----TPPALR----- 822
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 771 IETPDLTLHPIELIARQLTLESPLYNAVOPSELVGSVMTKEDKEINSPLLMKIMHTN 830

```

```

OY 823 -----EASELRELCMKNNMYKCYCSI--SPEMK----- 849
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 831 LTLWFEKCIYETENLEERNAVVRITIELIQLVQELNNNGVLEVVASAMSSPVYRLDHTF 890

OY 850 QQL--QOKAVASEIFKGNKY-----POSVP--RLFISTRGTEISPRVLQS 894
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 891 EQIPSRQKKILIEAHEISEDHKKYKLAKRSINPCVPFEGYLTNIIKTEEGNPEVLKR 950

OY 895 LGSEPIQYA-----VPVYKDYDRKGRPRQOLLTPSAVVIYVDADAKYKQRIYANLTGISV 950
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 951 HGKELINFSKRRKVAETGTGTQYQONOPYCLRV-----ESDIKREFENLDMGSM 1001

OY 951 -SSLSDSLFLVHQREDNOK 970
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1002 EKEFTDYLFRKSLIEPRNPK 1022

```

RESULT 15

```

US-08-290-731C-2
Sequence 2, Application US/08290731C
Patent No. 5843646
GENERAL INFORMATION:
APPLICANT: BOWTELL, David Douglas Lawrence
TITLE OF INVENTION: DNA MOLECULES ENCODING MURINE
TITLE OF INVENTION: SON OF SEVENLESS (MSOS) GENE,
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MTON, ZINN, MACPHEAK & SEAS
STREET: 2100 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,731C
FILING DATE: 17-OCT-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/AU93/00068
FILING DATE: 17-FEB-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PL0921/92
FILING DATE: 17-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: Q-36066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-731C-2

```

Query Match 2.38; Score 122.5; DB 2; Length 1319;
 Best Local Similarity 18.5%; Pred. No. 0.045;
 Matches 217; Conservative 170; Mismatches 359; Indels 425; Gaps 63;

```

OY 70 PYRDLQIYRQHMERKRGVSFYEPPLFLAVADTVYRALRTERDQAVMISGSGAGKTE 129
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 PY---EFFSEBNAPKRWGL---LVPALKKVOGQVHPTL--ESNDALQYVEE----- 50

```

QY 130 ATKRLQFYAETCPA-DEKGAVERDILQ--NEVLE-ATGNAKTLINDNSRREGKMYD 185
 Db 51 ---LILQILNMLCOQPRASADVERVOKSFPHIDKMAIDAOASALEKRRR----- 100
 QY 186 QFDRKGAQVGHILSTYLLKSRVYHONGHGERNFVFO-----LLEGEETIRRLGL 238
 Db 101 -----NPLS--LPAERIHILLREVLYGKIDQVSVYIYAVLEYISADILKLVG- 146
 QY 239 ERNPOSYL-----YLKGOCAKVSINDKSDWKVRRKALSVITFTEDEVEDLSTVASVL 293
 Db 147 -----NYVRNIRHYEITRKODIKVAMCADK-----VLMDFHODVEDI----- 183
 QY 294 HLGNIHRADED-----SNAOVTEN 314
 Db 184 ---NILSLTDEPSTSGQIYYDLVKAFMAEIROYLRELNLIIVKREPEVYNSKLFSSN 240
 QY 315 QUKYL-----FRLG-VEGTT--LREALTHRKIIAKGELLSPMLN-0AAVAR 359
 Db 241 DVENIFSRIVDIHELSTVLLGHIEDTYEMTDEGSPHPLVSGCFEDLAEELAFDYESYAR 300
 QY 360 DAL-----AKAVYSRT-----FTWLVRKI 378
 Db 301 DILRFGFHGHEFLSOLSKPGALYLQSIGEFKEAVOYVLPRLILAPYHCLHYFELLKOL 360
 QY 379 NRSLASKDAESPMSWSTVLGLDIY--GEVEFOHNSF-----EFCINYNCE-KLOOLF 430
 Db 361 BEK--SEDOEDKECKKQAITALLANQSGMEKICSLAKRRLSESACRFYSQOMKQOLA 418
 QY 431 IELTLKSEOEYEAGIAMEPVQYFNKIID--LVEKFKGIISILD--EBCLR--- 481
 Db 419 IK--KMEIOKNIIDGMEGKDIGOCNEFIMEGTLTVGAKHERHIFLEDGIMCCKSNH 475
 QY 482 ---PGFAT-----DLFLEKLEDIVKPHPHFLTHKLADQKTRKSLDRGEFR---L 525
 Db 476 GQRLPGASSAYRIKEKPFMRKVQINDK-----DTSEYKHAFFII 517
 QY 526 LHYAGEVYVYTGFLDKNN-----DLFRNLKETMCSMNPIMAQCFKSELSDKRPE 579
 Db 518 LKDGNSVIFSAKSAEKKNNMMWMAALISLOYSTLERM-----LDVTYLOEKE- 564
 QY 580 TVATQFKMSLLOVELLSKREPAYTRCIKPNDAQGRFDE-----VLTRHOQVYGL 632
 Db 565 -----EOWRLPSAEYRFAEP-----DSEENILFEENWQKAGIPITKAGTVLKL 609
 QY 633 MENLVRBRAGFAYRRKYEAFLORYKSLCETPMWAGRPDQGAVALVRLHLY--KPEYKM 691
 Db 610 IERLITYHM--YADPNVFRFTLYRSFC-----RQELLSLIERFEIPEPEPTEA 658
 QY 692 GRTKIFIRPKTLFATED-----SLEVR-----QSLATKIQAMRGFHW--ROKFLRYK 739
 Db 659 DR-----IAIENGDOPLSALKRFRREYIOPVOLRLVNCR--HWVEHHFYDFE 705
 QY 740 RSLICTOSW--WRGTGRRKAKRKMAOTIRRLR-----GFIHRSPRCPE 785
 Db 706 RQADLLOQMEEFITGY--RGKAMKKV--VESTIKIIQKKIARDNGRGHNTFQSSP--PT 760
 QY 786 NAEFLD---HVRASFLNLR-----ROL-----PRNVLDTSWP-----TPPP 819
 Db 761 VEMHISRPHIEFFDLTLHPLEIAQULLESPLYRAVOPSELVGSVTKEDKEINSPN 820
 QY 820 ALR-----EASELLRELCMKMVMKYCRSI-- 844
 Db 821 LTKMIRHTNLTLMFEKCIETENTLEERAVVSRILEILOVFOELNNFNGVLEVVSAAMS 880
 QY 845 SPEWK-----OOL--OQKAVASEIFKGRKDY-----POSVP--RLFISTRIGT 884
 Db 881 SPVYRLDHTFEQPSQOKKILBERAHLSHDHYKYLAKRLSINPCVPPFGIYLTNILKT 940
 QY 885 EETSPRYLOISGEPIQYA---VPVYKYDRKGYKRPQOLLITPSAAVYVEDAKYKQRI 940
 Db 941 EECNPEVLRHKGELINFSKRRRVALETGEIQQYQNOQPCYCLRVEPD-----IKRFF 991
 QY 941 DYANLIGISV--SSLSDSLFVLHVQREDNKK 970

Db 992 ENLNPWGNMKEKEFTDYLFNKSLIEPRHDK 1022

Search completed: July 7, 2003, 14:25:39
 Job time : 43.4566 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OH protein - protein search, using sw model

Run on: July 7, 2003, 14:21:30 ; Search time 43.3358 Seconds

(without alignments)
2770.284 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398

Sequence: 1 MRYRASALGSDGVRTWESA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppaa/PCT08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1285.5	23.8	2058	9	US-09-815-379-17
2	1261.5	23.4	1016	12	US-10-044-303-2
3	1241.5	23.0	2048	9	US-09-815-379-12
4	1241.5	23.0	2057	9	US-09-815-379-10
5	1240	22.1	765	12	US-10-044-303-1
6	1192	22.1	1938	9	US-10-171-311-164
7	1192	22.1	1972	9	US-10-171-311-162
8	1188.5	22.0	1945	9	US-09-927-597-2
9	1188.5	22.0	1979	9	US-09-927-597-4
10	1186	22.0	569	10	US-09-925-300-1583
11	1175.5	21.8	845	9	US-09-927-597-14
12	1172	21.7	849	9	US-09-927-597-12
13	1172	21.7	852	9	US-09-927-597-10
14	1168.5	21.6	544	9	US-10-106-698-5864
15	1150.5	21.3	1285	9	US-10-205-823-273
16	1132	21.0	2099	9	US-10-128-714-3290
17	1130	20.9	2405	9	US-10-128-714-8290
18	1127	20.9	772	9	US-09-927-597-8
19	1105	20.5	2548	10	US-09-851-682A-1

20	1042	19.3	699	9	US-09-927-597-6	Sequence 6, Appl1
21	568	10.5	245	10	US-09-925-302-686	Sequence 686, App
22	566	10.5	570	10	US-09-764-853-551	Sequence 551, App
23	497	9.2	2568	10	US-09-866-108-3	Sequence 3, Appl1
24	244.5	4.5	294	9	US-10-106-698-5501	Sequence 5501, App
25	215.5	4.0	286	10	US-09-925-301-923	Sequence 923, App
26	207.5	3.8	237	9	US-09-925-299-937	Sequence 937, App
27	207.5	3.8	237	10	US-09-925-299-937	Sequence 937, App
28	160	3.0	2549	9	US-09-950-634-3	Sequence 3, Appl1
29	150.5	2.8	103	10	US-09-864-761-43208	Sequence 34208, A
30	150.5	2.8	103	10	US-09-864-761-41932	Sequence 41932, A
31	148	2.7	77	10	US-09-864-761-48341	Sequence 48341, A
32	145.5	2.7	1337	10	US-09-803-146-1	Sequence 1, Appl1
33	144.5	2.7	100	10	US-09-864-761-39854	Sequence 39854, A
34	144.5	2.7	102	10	US-09-864-761-41762	Sequence 41762, A
35	144.5	2.7	134	9	US-10-091-438-227	Sequence 227, App
36	144.5	2.7	134	10	US-09-764-853-760	Sequence 760, App
37	140	2.6	89	10	US-09-864-761-41891	Sequence 41891, A
38	137.5	2.5	103	10	US-09-864-761-42604	Sequence 42604, A
39	136.5	2.5	103	10	US-09-864-761-42606	Sequence 42606, A
40	133.5	2.5	1045	10	US-09-815-242-10617	Sequence 10617, A
41	132.5	2.5	3899	9	US-10-171-311-4	Sequence 4, Appl1
42	132.5	2.5	3907	9	US-10-171-311-2	Sequence 2, Appl1
43	132.5	2.5	3917	9	US-10-171-311-8	Sequence 8, Appl1
44	132.5	2.5	3925	9	US-10-171-311-6	Sequence 6, Appl1
45	130.5	2.4	66	10	US-09-864-761-42111	Sequence 42111, A

ALIGNMENTS

RESULT 1
US-09-815-379-17
Sequence 17, Application US/09815379
Publication No. US20030073613A1
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA
APPLICANT: GERITSEN, MARY
TITLE OF INVENTION: ANGIOGENESIS ASSOCIATED PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE: 10716/35
CURRENT APPLICATION NUMBER: US/09/815, 379
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 60/191,134
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 2058
TYPE: PRT
ORGANISM: Homo sapiens
US-09-815-379-17

Query Match 23.8%; Score 1285.5; DB 9; Length 2058;
Best Local Similarity 36.9%; Pred. No. 5.6e-90;
Matches 301; Conservative 150; Mismatches 250; Indels 115; Gaps 20;
QY 3 YRASALGSDGVRTWESA...LTARDKGVODFVLENTFSEANFIENLRRENTIYYIG 62
DB 44 YKOSTTHOKV-----TAMPTNEGVDMASTLELHG-GSIMYNLRFQRYKRNQIYYIG 97
QY 63 PVLVSVPYRDLC-IYSRQMERRYRGVSFEVPHPLFAVDYVYRALRTERPDQAVWSIG 121
DB 98 SILASVNPYQPIALYEPATMEQYSRHLGELPRHIFAIINECRCLMKRYDNCILISG 157
QY 122 ESGGKTEAKTRKLLQFAVE-----TPADRGCAVDRLQSNPVLAEFGNAKTYNNN 175
DB 158 ESGGKTESKRLIKFLSVISQSLSELEKTESCAVRAILESSPIEAFGNKATYNNN 217
QY 176 SSRFGKYMDVQFDEKGAIPVGSHITSYLEKSRVYHONGERNRHFVQLLEGDEEFLTR 235
DB 218 SSRGKTVQNLICQKGNIGQGRIVDTLEKRNRYVRPNGBERNHIFALLAGLEHEREE 277

; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 2048
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-815-379-12

Query Match 23.08; Score 1241.5; DB 9; Length 2048;
 Best Local Similarity 36.1%; Pred. No. 1,4e-86;
 Matches 295; Conservative 154; Mismatches 252; Indels 117; Gaps 22;

```

OY 3 YRASALGSDGVRYTMSALTAARDVGVODFVLLNFTSEAFLENTLRARRRENTLYTYIG 62
DB 41 YKOSTITHOKV-----TAMHPTEBEGVDDMASTLELHG-GSIMYNLFQKRNQIMTYIG 94
OY 63 PVLVSVPYRDLO-TYRQHMERYRGVSFEYEPHLEFAVDYVYALRTERDOAVMISG 121
DB 95 SILASVNPYQPIAGLEVAPATMEQYSRHHGELRPHIFALANCYCRLMRHNOICILKG 154
OY 122 ESGAGTEATKRLLOFYAE-----TCPAERGAARBDLLOSNPYLEAFNAKTLRDN 175
DB 155 ESGAGTEATKRLKLFSLVISOQSLSELSKECTSCVERAILLESSPIMEAFNAKTYVNN 214
OY 176 SSRGKYMDOFDFKAPVGGHILSTYL-EKSRVYHONHGERNFHFVYOLLEGGEEETLR 234
DB 215 SSRGKRVQNLQICQKNGIOGRIVDCILSSQNRVYRQNGERNYHIFALLAGLEHERE 274
OY 235 RLGIERNPOSYLYLVGQCAKVSINDSKDMYMRKALSVIDFTEDEVEDLLSTYASVLH 294
DB 275 EFLYS-TPENYHNLNOSGVEDKTIISDQSFREVITAMDVMOFSKEEVREYSRLLAGILH 333
OY 295 LGNIHFADEDSNAOYTTEENOLKYLRLIGVEGTLREALTNRKILIAEGEELSPLENO 354
DB 334 LGNIEFT--AGGAQVSFTALGRSAELGLDPTQLTALQORSFLEGEELTPLVNOQ 391
OY 355 AAYARDALAKAVYSRTFTVLVRKIN-RSLASKADSPSRSTTVYGLLDYGFVEYQHS 413
DB 392 AYDSRSLMALALACCFEYVIRKINSRIKGNEDFS-----IGLIDIGFENFEYNH 443
OY 414 FEOFCINYNCKLQOQLFELTLKSEOEYEAEGIMEPVQYFNKIKIDLVEEKFKIIS 473
DB 444 FEOFNINYNANKEKLOEYFNHIFSLQLEYSRGLWEMEDIDWNECIDLLEKTL-GLLA 502
OY 474 ILDEBCLRGAEATDLFLEKLEDYVPRPHFLTHKLADQKTRKSLDRGEFRLHYAGEVT 533
DB 503 LINESHHP-QATDSTLEKLH-SQHANNHYY-----KPRVAVN--NGCVKHYAGEVQ 552
OY 534 YSVTGFELDKNNDLLEFNLKETWCSMNPIMAOCE-----DKSELSDKKRPETVATQ 584
DB 553 YDVRGILEKNRDTFDDLLNLRSERFIDYLEHVSRRNODTLCKGSHRRPTVSSQ 612
OY 585 FKM-SILOLVELILRSKEPAYIRICIRPNDAKOPGREDEVILRHQVYGLMELNLRVAGF 643
DB 613 FKVDSLHSLMATLSSNPFYRCIRPNQKMPDDQAVVNLQRLYSMLLETVAIRKAGY 672
OY 644 AYRKRYEAFLOKYSCLPETPMNAGRPQDQAVVNLHLYGKPE-----687
DB 673 AYRRPPODFYKRYK-----VLMRNAL--PEDVAGKCTSLLOLYDAS 712
OY 688 --EYKMGRTKLFIREPKTLFATEDSLEVRROS-----LATK-----721
DB 713 NSEMOLGKTKVFLR-----ESLEQLEKREEREESHAMVYIRAHVIGFLARKQYRVLYC 767
OY 722 ----IAAMRGFHMROKFLVRKRSALCIDSMMRGTLGRR 756
DB 768 VVIQKNIRAFLLRRRLHLKKAIVFOKOLRGQIARR 805
  
```

RESULT 4
 US-09-815-379-10
 ; Sequence 10, Application US/09815379
 ; Publication No. US20030073613A1

; GENERAL INFORMATION:
 ; APPLICANT: RASTELLI, LUCA
 ; APPLICANT: GERRITSEN, MARY
 ; TITLE OF INVENTION: ANGIOGENESIS ASSOCIATED PROTEINS AND NUCLEIC ACIDS
 ; TITLE OF INVENTION: ENCODING THE SAME
 ; FILE REFERENCE: 10716/35
 ; CURRENT APPLICATION NUMBER: US/09/815,379
 ; CURRENT FILING DATE: 2001-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,134
 ; PRIOR FILING DATE: 2000-03-22
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 2057
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-815-379-10

Query Match 23.08; Score 1241.5; DB 9; Length 2057;
 Best Local Similarity 36.1%; Pred. No. 1,4e-86;
 Matches 295; Conservative 154; Mismatches 252; Indels 117; Gaps 22;

```

OY 3 YRASALGSDGVRYTMSALTAARDVGVODFVLLNFTSEAFLENTLRARRRENTLYTYIG 62
DB 41 YKOSTITHOKV-----TAMHPTEBEGVDDMASTLELHG-GSIMYNLFQKRNQIMTYIG 94
OY 63 PVLVSVPYRDLO-TYRQHMERYRGVSFEYEPHLEFAVDYVYALRTERDOAVMISG 121
DB 95 SILASVNPYQPIAGLEVAPATMEQYSRHHGELRPHIFALANCYCRLMRHNOICILKG 154
OY 122 ESGAGTEATKRLLOFYAE-----TCPAERGAARBDLLOSNPYLEAFNAKTLRDN 175
DB 155 ESGAGTEATKRLKLFSLVISOQSLSELSKECTSCVERAILLESSPIMEAFNAKTYVNN 214
OY 176 SSRGKYMDOFDFKAPVGGHILSTYL-EKSRVYHONHGERNFHFVYOLLEGGEEETLR 234
DB 215 SSRGKRVQNLQICQKNGIOGRIVDCILSSQNRVYRQNGERNYHIFALLAGLEHERE 274
OY 235 RLGIERNPOSYLYLVGQCAKVSINDSKDMYMRKALSVIDFTEDEVEDLLSTYASVLH 294
DB 275 EFLYS-TPENYHNLNOSGVEDKTIISDQSFREVITAMDVMOFSKEEVREYSRLLAGILH 333
OY 295 LGNIHFADEDSNAOYTTEENOLKYLRLIGVEGTLREALTNRKILIAEGEELSPLENO 354
DB 334 LGNIEFT--AGGAQVSFTALGRSAELGLDPTQLTALQORSFLEGEELTPLVNOQ 391
OY 355 AAYARDALAKAVYSRTFTVLVRKIN-RSLASKADSPSRSTTVYGLLDYGFVEYQHS 413
DB 392 AYDSRSLMALALACCFEYVIRKINSRIKGNEDFS-----IGLIDIGFENFEYNH 443
OY 414 FEOFCINYNCKLQOQLFELTLKSEOEYEAEGIMEPVQYFNKIKIDLVEEKFKIIS 473
DB 444 FEOFNINYNANKEKLOEYFNHIFSLQLEYSRGLWEMEDIDWNECIDLLEKTL-GLLA 502
OY 474 ILDEBCLRGAEATDLFLEKLEDYVPRPHFLTHKLADQKTRKSLDRGEFRLHYAGEVT 533
DB 503 LINESHHP-QATDSTLEKLH-SQHANNHYY-----KPRVAVN--NGCVKHYAGEVQ 552
OY 534 YSVTGFELDKNNDLLEFNLKETWCSMNPIMAOCE-----DKSELSDKKRPETVATQ 584
DB 553 YDVRGILEKNRDTFDDLLNLRSERFIDYLEHVSRRNODTLCKGSHRRPTVSSQ 612
OY 585 FKM-SILOLVELILRSKEPAYIRICIRPNDAKOPGREDEVILRHQVYGLMELNLRVAGF 643
DB 613 FKVDSLHSLMATLSSNPFYRCIRPNQKMPDDQAVVNLQRLYSMLLETVAIRKAGY 672
OY 644 AYRKRYEAFLOKYSCLPETPMNAGRPQDQAVVNLHLYGKPE-----687
DB 673 AYRRPPODFYKRYK-----VLMRNAL--PEDVAGKCTSLLOLYDAS 712
OY 688 --EYKMGRTKLFIREPKTLFATEDSLEVRROS-----LATK-----721
DB 713 NSEMOLGKTKVFLR-----ESLEQLEKREEREESHAMVYIRAHVIGFLARKQYRVLYC 767
  
```

722 ---10AAMRGFMHROKELRYKSAICISQMSWGTLGR 756
768 VVIOKMYRAFLRRRLHLKKAIVFOKOLRGQIARR 805

RESULT 5
US-10-044-303-1

Sequence 1, Application US/10044303
Patent No. US20020137161A1
GENERAL INFORMATION:
APPLICANT: Max-Planck-Gesellschaft e.V.
TITLE OF INVENTION: Protein expression and structure solution using
FILE OF INVENTION: Specific fusion vectors
FILE REFERENCE: ST010209-EPA
CURRENT APPLICATION NUMBER: US/10/044.303
CURRENT FILING DATE: 2002-01-11
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentlin Ver. 2.1
LENGTH: 765
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Partial myosin
OTHER INFORMATION: sequence of Dictyostelium: Component (1) of the
OTHER INFORMATION: recombinant protein M761-2R R238E
US-10-044-303-1

Query Match 23.0% Score 1240; DB 12; Length 765;
Best Local Similarity 38.1% Pred. No. 4e-87;
Matches 265; Conservative 154; Mismatches 248; Indels 28; Gaps 12;

28 GVQDFVLENTSEAFATLNRFRRENLITTYIGPVLSVNPYRDQIYSROHMERYRG 87
91 GVEDMSEL-STYNEPAVPHNLRYRNOGLIYSGFLVAVNPFRIPITQEWVDFKG 149
88 VSEYEPHPLFAVATYVTRALRTERRDQAVMISGSGAGTEARRKLLQFAECPAPER 147
150 RRNEVAPIFAISVAPRSMDDKONOSLITGSGAGTEARRKLLQFAECPAPER 209
148 GGA--VRDLQSNVLEAFNAKTLRNDNSRFQYMDVDFGAPVGHILSYLLEK 205
210 NSGVLEQOILQANPLLEAFNAKTLRNNNSRFQYMDVDFGAPVGHILSYLLEK 269
206 SRVYHONGERNFHYFQQLLEGEETLRRLGLERNPOSTYLIVKGCQAKYSINDSDW 265
270 SRVYHONGERNFHYFQQLLEGEETLRRLGLERNPOSTYLIVKGCQAKYSINDSDW 328
266 KVMKALSYIDFTEDEVEDLLSIVASVLIHGNHFAADEDSNAOVTTEENOLKYLRLIG 325
329 KIRQAMDIVGSGEOMKIFKILGILHGNHFAADEDSNAOVTTEENOLKYLRLIG 388
326 EGTTLREALTHRKIIAKGELLSPNLQAAVAPALAKAYISRTFTLVKIRNSLASK 385
389 NPVLEKALMEPRILLAGRDVAQHINVERSSSRDALVKAALYGRFLMLVKKINNVCQ 448
386 DASPBRSTTVYGLDIFYGEVFOHNSFEQCIYCNKELCOULFIELLSQOEYEA 445
449 -----RAYFEGVLDISGEFELFKVNSFQICINTNEKLOQFNHMHFLEQETLKE 501
446 GIAMEPVQY-FNNKIIICDVE-EKFGKIIISILDECLRGEATDLPLEKLEDTYVKKPH 503
502 KIMWTFIDGLDSQATIDLDGQPGIALLDEQSVFP-NATDWTLLTKL-----HSH 554
504 FLTHKLADOKTRKSLDRGEFRLIHYAGEVTVSYTGFLDKNNDLRLNKEKMSMNP 563
555 FSKRNKAYEPRPS--KTEGVYHAGVAYEIQWLEKKNKPLDQDDLECFKSSDVAV 612
564 AOCFDSKSLDKKRP-----TVATQKMSLQIIVELIRSKEPAYIRKICIPNDAKOPGRFD 619
613 TKLENDPNINASRAKKGANFTTVAAYQKQELASIMATLETTNHFVRCIIPNNKOLPAKLE 672

620 EVLIRHGVYKILGIMENLRVRAAGFAYRKYEAFLQKSLCPETWPMNAGRPQDQAVLV 679
673 DKVYDOLRONGVLEGIRITRKGPNNKIYADFKYIILANV-PROAEDSQKATDAVL 731
680 RHLGKPEEKYMGKTKIFIRFPKTLFATEDSLSEVR 714
732 KHLNIDPEQYRFGTITKIFFRAGQ-LARIEAFRROR 765

RESULT 6
US-10-171-311-164

Sequence 164, Application US/10171311
Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumel
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangl
APPLICANT: Glati, Karen
APPLICANT: Gannavarapu, Manjula
APPLICANT: Hoersht, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
FILE OF INVENTION: OF CERVICAL CANCER
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171.311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
PRIOR FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FASTSEQ for Windows Version 4.0
LENGTH: 1938
TYPE: PRT
ORGANISM: Homo sapiens
US-10-171-311-164

Query Match 22.1% Score 1192; DB 9; Length 1938;
Best Local Similarity 29.5% Pred. No. 8.7e-83;
Matches 316; Conservative 197; Mismatches 357; Indels 200; Gaps 30;

40 SEAAFTENLRFRRENLITTYIGPVLSVNPYRDQIYSROHMERYGVSEYEPHPLFA 99
97 NEASVLIHNRERYFGLIYTYSGFLCVVNPYKHLPIYSEKIVDMYKGRKHEMPHIA 156
100 VADTVTRALRTERRDQAVMISGSGAGTEARRKLLQFAECPAPERG-----GAVR 152
157 IADTAARSMLODREDSILCTGESAGKTEARRKLLQFAECPAPERG-----GAVR 215
153 DRLLQSNVLEAFNAKTLRNDNSRFQYMDVDFGAPVGHILSYLLEKSRVYHON 212
216 KOLLQNPILLEAFNAKTLRNDNSRFQYMDVDFGAPVGHILSYLLEKSRVYHON 275
213 HGERNHVYQQLLEGEETLRRLGLERNPOSTYLIVKGCQAKYSINDSDWVMKAL 272
276 RDBRTFHIFVYAGAKEMKMRSDLLLE-GFNNTYPLISG-FVPIPAADDDMEFOETVEM 333
273 SVYDFTEDEVEDLLSIVASVLIHGNHFAADEDSN-AOVTTEENOLKYLRLIGEGTTL- 330
334 AIMGFEERQSLTKVAVSSVLQGNIVRKKEKNTQASMPDNTAQAQKCHLMGINVDF 393
331 REALTHRKIIAKGELLSPNLQAAVAPALAKAYISRTFTLVKIRNSLASKDAES 389
394 RSLITFR--TKGROVYQAQKREQADFAVEALAKATYERLFRWILFRVNNALDKTRHOG 451
390 PSNRSTTVGLDIFYGEVFOHNSFEQCIYCNKELCOULFIELLSQOEYEAEGTAV 449
452 ASF-----IGIIDIAGFEIPEVNSFQICINTNEKLOQFNHMHFLEQETLKE 506

[illegible][illegible]

```

Db      216  FAYGELKOLLQAPILFAGNMAKTVANNSSRFGRFIRNFPVITYIGANITETLLK 275
Qy      206  SRVYHONHGERNFVIFQOLLEGEEELRLGLERNPOSLYLVKQCAKVSINDKSDM 265
Db      276  SRATQRARDEFTFIEFYMIAGAEKMRSDLEL-GEFNVTFLSNG-FVPIPAQODEMF 333
Qy      266  KVMKRALVIDFTEDEVEDLSIVASVHLGNHFAADDSN-AQVTTENQAKYLRLLG 324
Db      334  QETVEAMVINGFSEEDLSILKYVSSVLQGNIVFKERNTDQASHPDPTAAQKCHLMG 393
Qy      325  VEGTTL-REALTNRKTIAGEELLSPILN-EQAAYADALAAVSRFTWLVKRNLSL 382
Db      394  INWTDFTRSILTPR--IKVGRDYVQAQKTEQADFAVEALAKATERLFRWTLITVKNAL 451
Qy      383  ASKDAESWRSTTVGLDIYGEVFOHNSFPQFCINCNKLOLFELTLKSEOEY 442
Db      452  DKHROGASF-----LGLIDLAGELFEVNSFEQICINTYNEKLOLFNHTFLEOEY 506
Qy      443  EAGGIAMEPVQYFNKNIIC-DIVEERK--GIISILDEBCLRPGEATDLFLEKEDTVK 499
Db      507  QREGIENWFIDGLDQPCIELIERPNPNPGVALLDEDCMPF-KATDKSFYEKLTGEG 565
Qy      500  PHHFLPLHLAODKTKSLDRGEFRLHAGETVTVSGFLDKNNDLFRNLKETMCSM 559
Db      566  SHKF-----QKPKQDKTEFESIHYAGKYVNASAMLTKNMDPLNDVYLSLNAS 618
Qy      560  NPIMAOCF-----DKSELSDKRP-----EYATQFKSLQVLEILASK 599
Db      619  DKFVADLMKDVDRIVGLDQMAKTESLPSASKTKKGRFVYQOLYKOLKMLTTLNT 678
Qy      600  ERYVITICIRPNDAKQGRFDEVLIHQVYKYLGLMENVLRVRAFGAYRRKYEAFLORYKSL 659
Db      679  TPNFVRCIIPNHEKRSGLDAFLVLEQLRCNGVLEGRICRQGFNRNIVPQEFQRYEIL 738
Qy      660  CPETVPMWAGRPQDGVAVLVRHLGKPREYKMGKRIKITRPKTLFATEDSLVYRROSLA 719
Db      739  AANAIRKGMQKOCILMIKALELDPLNLYRIGOSKIFFR-TGVLAHLEBERDLKIDVI 797
Qy      720  TKIQAAAMRGFHRKOKFLRVKRSALCIO-----SMWR----- 750
Db      798  MAFQAWCRGYLARKKAFKAKQOOLUAMKVIQRNCAVYLKLRWQWHRFLTKKPLLOYTRQ 857
Qy      751  -----CTLGRRAAKRKAQAOTIRRLINGEILIRHSPCPENAFELDHVASFLLNR 802
Db      858  EEMQAKEDLOKTEROOK-AENELKELEQ-----KHSQLTBE----- 895
Qy      803  ROLPRNVLDTSWTPPALREASELRELCKMNMWYKCRSISPEMGOLOQKAVASEIF 862
Db      896  -----KNLQEOLOQATELXAEHEKRVYLAAK-----KQELKEE--ILHE-- 933
Qy      863  KKKKNDYPOSVPRLEFISTRIGTEISPRVQSLGSEPIQVAVPVVYKDRGKYKRPQOL 922
Db      934  -----MEARLEEEEDRGQLOA-----ERK--KMAQOMLD 961
Qy      923  LTPSAVVIYEDAKVQRDVANLNG-LSVSSLSLFLVLAHQREDNQKGDVYQOSHTV 961
Db      964  LEEQ---LEEBEAAQKIQLEKVTAAEAKIKKLEDELLVMDQ--NNLSKERKLEERIS 1016
Qy      962  ETLVTALSDRVNNININOSITFAGPGRGDIIDFTSGSELLITKAKNGHAAVYA 1038
Db      1017  DLTWMAEEEEKAKN-----LTKLKKHSHSMIS 1044

```

RESULT 10

```

; Sequence 1583, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101

```

```

; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1583
; LENGTH: 569
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (291)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (345)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (346)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (552)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (553)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (554)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-300-1583

Query Match      22.0%; Score 1186; DB 10; Length 569;
Best Local Similarity 47.1%; Pred. No. 3,8e-83;
Matches 243; Conservative 86; Mismatches 167; Indels 20; Gaps 5;

Qy      285  LLSIVASVHLGNHFAADDSN---AQVTTENQAKYLRLLGVEGTLREALTHRKII 340
Db      28  VLAAYAVYLVKLTGNIIEFKPESRYNGLDESKIKDKNELKICELTIDDSVLEAFSFTVE 87
Qy      341  AKGEELSPILNEQAAYARDAALAAVSRFTWLVKRNLSLASKDAESPSWRSTVLGL 400
Db      88  AKQEVSTVLNVAQATYARDALAKNLSRLFSWLVNRINSEIKQOTVRRK-----VMGV 142
Qy      401  LDYGEVFOHNSFPQFCINCNKLOLFELTLKSEOEYEAEGIAMEPVQYFNKNI 460
Db      143  LDYGEFELFEDNSFPQFCINCNKLOLFELTLKSEOEYIREXLEWHIDFNNALI 202
Qy      461  CDVLEERKGIISILDEBCLRPGEATDLFLEKEDTVKPHHFLTH--KLADQKTRKSL 518
Db      203  CDLIENNTNGILAMIDEBCLRPGEATDLFLEKEDTVKPHHFLTH--KLADQKTRKSL 262
Qy      519  DRGEFRLHAGETVTVSGFLDKNNDLFRNLKETMCSMNPIMAOCPFKSELS--DKK 576
Db      263  PHSCFRIGHYAGKVLYQVEGVVDKNDLXTRDLSQAMWKASHALINSLEPEGNPATINK 322
Qy      577  RPEYVATQFKKSLQVLEILRSKREPAYIRICIKPDNDKQGRFDEVLIHQVYKYLGLMENV 636
Db      323  RPTAGSQFKASVATLKKNIQTXXPNYIRICIKPDNDKKAHIFDALVYCHQIRVLGLLENV 382
Qy      637  RVRAGFAVRRKYEAFLORYKSLCPETVPMWAGRPQDGVAVLVRHLGKYRDEYKMGRTYI 696
Db      383  RVRAGFAVRRKYEAFLORYKSLCPETVPMWAGRPQDGVAVLVRHLGKYRDEYKMGRTYI 442

```



```

Db 507 QREGIENFIDFGDLDPICIELERPNPPGVALLLDEBCEWFP-KATDKSVELELCTEG 565
Qy 500 PHPEFLTHKLADQTRKSLDGRGRLHAGEVYVSTGFLDKNDLLFRNLKETMCSM 559
Db 566 SHKF-----QKPKLKDTEFSIIHYAGKVDYNASAWLTKMNDPLNNTSLMSS 618
Qy 560 NPIMACF-----DKSELSDKRRP-----EVATOFKMSLLOVELIRSK 599
Db 619 DKFVADLMKDVRIVGLDMAMKTESLPSASKTKKGMRTYGOULYKELGKMTTLTNT 678
Qy 600 EPAYICIKPNDAKOPRDEVLIRHOVYLGMLNLRVRAGFAVRRKYEAFLQRYKSL 659
Db 679 TPWFVNCIIPNHEKRSGLDAFLVLEOLRCNGVLEGIRICROGFPNRIYFOEFRQREYL 738
Qy 660 CETWPMWAGRPDQGVAVLVRLHGYKPEYKMGRTKFIREFPKTLFATEDSLVROSLA 719
Db 739 AANAIPKGFMDGKACILMTKALELDPNLYRIGOSKIFFR-TGVLAHLEERDLKITDVI 797
Qy 720 TKIOAMRGFHWROKFLRVKRSACIO-----SMWR 750
Db 798 MAFOAMCRGYLARKAFKROOQLTAMKVIOIRNCAAYLKLNNQWNR 843

```

```

RESULT 13
US-09-927-597-12
/ Sequence 12, Application US/09927597
/ Publication No. US20030032018A1
/ GENERAL INFORMATION:
/ APPLICANT: Malik, Fady
/ APPLICANT: Beraud, Christophe
/ APPLICANT: Freedman, Richard
/ APPLICANT: Craven, Andrew
/ APPLICANT: Sakowicz, Roman
/ APPLICANT: Hartman, James
/ TITLE OF INVENTION: Human smooth muscle myosin heavy chain
/ FILE REFERENCE: CYTOPI8
/ CURRENT APPLICATION NUMBER: US/09/927,597
/ NUMBER OF SEQ ID NOS: 17
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ LENGTH: 852
/ TYPE: PRT
/ ORGANISM: Human
US-09-927-597-12

```

```

Query Match 21.7% Score 1172; DB 9; Length 852;
Best Local Similarity 35.0% Pred. No. 8.5e-82;
Matches 268; Conservative 157; Mismatches 267; Indels 74; Gaps 17;

Qy 40 SEAAFIENLRRERENLIYTYIGPVLYSVNPRDQIYSGRQHMERYRGVFEVPHLFA 99
Db 97 NEASVLAHNRERFSLIYTSGLPCVVVNPYKHLPIYSEKLYDMKKGKREMPHITA 156
Qy 100 VADTVYRALRTERDQAVMISGSGAGKTEATKRLQYAEPCPAERG----- 148
Db 157 IADTAIRSMLODRDOSIICGSESGAGKTEATKRYOYLA-VVASSHKGKDTSTIGPS 215
Qy 149 ---GAVRDLQSNPLLEAFNAKTLRNNDSSRFGMYOVODFGAPRGHIIISLEK 205
Db 216 FAYGELEKOLLQANPLEAFNAKTYKNDSSRFGFTRINDVGYIVGAVIETYLEK 275
Qy 206 SRVYHONHGERNFHFYQLLEGEDETLRLGLERNPOSYLYLVGOCAKVSSINDKSDM 265
Db 276 SRAIRQARDETFHFIYIMAGAKEMRSDLLLE-GFNNTPLMSG-FVPIPADDEMF 333
Qy 266 KVMRAKLSYIDTEDEVEDLSTIVSVHLGNIHRAADESN-AQVTTENOLKYLTRLIG 324
Db 334 QETVEAMATMGSEEQSLTKVSVSVQGLNIVFKKEKNDQASMPNTAAQKCHLMG 393
Qy 325 VEGTTL-REALTNRKIIAGBELLPLNL-EQAAVARDALAAVYSRTFTMLVLRINSL 382

```

```

Db 394 INTDTRSLIIFR--IKVGRDVAQAKQTEQADFAVEALAAATYERLEFRWILTRVKNAL 451
Qy 383 ASKDASPSWRSTTVGLLDIYGFVEFOHNSFEQFCINYNENKLOULFIELTKRQREY 442
Db 452 DKTHRGASP-----LGIIIDAGFELFEVNSPQDCINTNTKLOOLFHTMFIQEERY 506
Qy 443 EAGCIAMEYQVYFNKNTIC-DLYEEKF--GIISIDECLEARGENTDLPTEKLEDPVK 499
Db 507 QREGIENFIDFGDLDPICIELERPNPPGVALLLDEBCEWFP-KATDKSVELELCTEG 565
Qy 500 PHPEFLTHKLADQTRKSLDGRGRLHAGEVYVSTGFLDKNDLLFRNLKETMCSM 559
Db 566 SHKF-----QKPKLKDTEFSIIHYAGKVDYNASAWLTKMNDPLNNTSLMSS 618
Qy 560 NPIMACF-----DKSELSDKRRP-----EVATOFKMSLLOVELIRSK 599
Db 619 DKFVADLMKDVRIVGLDMAMKTESLPSASKTKKGMRTYGOULYKELGKMTTLTNT 678
Qy 600 EPAYICIKPNDAKOPRDEVLIRHOVYLGMLNLRVRAGFAVRRKYEAFLQRYKSL 659
Db 679 TPWFVNCIIPNHEKRSGLDAFLVLEOLRCNGVLEGIRICROGFPNRIYFOEFRQREYL 738
Qy 660 CETWPMWAGRPDQGVAVLVRLHGYKPEYKMGRTKFIREFPKTLFATEDSLVROSLA 719
Db 739 AANAIPKGFMDGKACILMTKALELDPNLYRIGOSKIFFR-TGVLAHLEERDLKITDVI 797
Qy 720 TKIOAMRGFHWROKFLRVKRSACIO-----SMWR 750
Db 798 MAFOAMCRGYLARKAFKROOQLTAMKVIOIRNCAAYLKLNNQWNR 843

```

```

RESULT 14
US-10-106-698-5864
/ Sequence 3864, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
/ FILE REFERENCE: PA0005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ PRIOR FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: PatentIn Ver. 3.0
/ LENGTH: 544
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (266)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (321)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (527)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (528)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (529)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5864

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:12:35 ; Search time 53.1849 Seconds

(without alignments)
1887.083 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398

Sequence: 1 MRYRASALGSDGVRTMESA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5261	97.5	1028	2	S37146 myosin I heavy chain
2	5143	95.3	1028	2	A59253 myosin I beta - hu
3	5128	95.0	1028	2	S41749 myosin heavy chain
4	4371	81.0	1028	2	I51173 myosin I beta - bu
5	4167	77.2	807	2	H75634 myosin-IC - mouse
6	2632	48.8	1026	2	S45574 myosin IB - fruit
7	2252	41.7	448	2	B45438 myosin I beta, MMI
8	2109	39.1	1078	2	B45439 myosin-I, Myr 1c (
9	2106.5	39.0	1043	1	A29483 myosin-I, Myr 1b (
10	2094.5	38.8	1107	2	C45439 myosin I alpha cha
11	2081.5	38.6	1079	2	A45438 myosin heavy chain
12	2080	38.5	1136	2	A45439 myosin heavy chain
13	2075.5	38.4	1094	2	S32404 myosin heavy chain
14	1904	35.3	1000	1	A33620 myosin heavy chain
15	1861	34.5	1006	1	A33933 myosin heavy chain
16	1803.5	33.4	1011	2	A55573 myosin IA - Caenor
17	1756	32.5	1017	2	T24349 myosin IA - mouse
18	1739	32.2	1003	2	S33760 myosin-IE heavy ch
19	1726.5	32.0	1099	2	A59300 myosin-IF heavy ch
20	1719.5	31.9	1099	1	S31926 myosin IB heavy ch
21	1715	31.8	1109	2	S53601 myosin-IC (similar
22	1712.5	31.7	1107	1	S52517 myosin I heavy cha
23	1692.5	31.4	1111	1	A33284 myosin heavy chain
24	1601	29.7	1113	1	A47106 myosin heavy chain
25	1588.5	29.4	994	1	J00151 myosin heavy chain
26	1567.5	29.0	1147	1	MMAX1B myosin heavy chain
27	1564	28.6	1100	2	A56511 myosin heavy chain
28	1545	28.6	1249	2	A56511 myosin heavy chain
29	1523.5	28.2	1168	1	MMAX1C myosin heavy chain

30	1509.5	28.0	1217	2	T39427	probable myosin I
31	1501	27.8	1219	2	S54570	probable membrane
32	1473	27.3	1215	2	T32734	myosin-IA - Acanth
33	1453	26.9	1181	2	T30578	myosin IC - slime
34	1450	26.9	1271	2	S37958	myosin heavy chain
35	1382.5	25.6	1446	2	T04528	myosin heavy chain
36	1361	25.2	2357	2	A59249	class VII unconvn
37	1354	25.1	1203	2	A59257	myosin VIIa, short
38	1354	25.1	1203	2	A59257	myosin VIIa, long
39	1352	25.0	1515	2	S51824	myosin heavy chain
40	1346	24.9	839	2	F83334	myosin heavy chain
41	1336	24.7	1528	2	T14279	myosin-like protei
42	1335	24.7	2121	2	A59233	myosin VII-like pr
43	1331	24.7	1477	2	T00957	myosin heavy chain
44	1331	24.7	2062	2	A59297	myosin X - mouse
45	1328	24.6	1611	2	A84743	probable myosin he

ALIGNMENTS

RESULT 1

S37146 myosin I heavy chain - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 02-Feb-2001

C:Accession: S37146

R:Ruppert, C.; Gogel, J.; Reinhard, J.; Baehler, M.

submitted to the EMBL Data Library, August 1993

A:Description: MYR-2, a novel class-I Myosin identified in rat brain.

A:Reference number: S37146

A:Accession: S37146

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1028 <RUP>

C:Superfamily: EMBL:X74800; NID:G400428; PION:CA52807.1; PID:G400429

C:Keywords: brush border myosin heavy chain I; myosin motor domain homology

F:14-683/Domain: myosin motor domain homology <MMOT>

F:105-112/Region: nucleotide-binding motif A (P-loop)

Query Match 97.5%; Score 5261; DB 2; Length 1028;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 1015; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY	17	MESALTARDVGVODFVLENTSEAFTEMLRRFRRENLITYIGPVYVNPYRDLQI	76
DB	1	MESALTARDVGVODFVLENTSEAFTEMLRRFRRENLITYIGPVYVNPYRDLQI	60
QY	77	YSRQHEHYRGVSFEVPPHFAVADTVYRALTERRRQDAVMIGESGAGTETKRLQ	136
DB	61	YTRQHEHYRGVSFEVPPHFAVADTVYRALTERRRQDAVMIGESGAGTETKRLQ	120
QY	137	FPAETCPAPERGAVERDRLQSNPYLEAFNAKTLRNDSSRFKQMDVDFGAPVGG	196
DB	121	FPAETCPAPERGAVERDRLQSNPYLEAFNAKTLRNDSSRFKQMDVDFGAPVGG	180
QY	197	HLSTLLEKSRVYHNGEERNPHVYQLLEGGEEETLRRLGLENNPOSYLYLVGQCAV	256
DB	181	HLSTLLEKSRVYHNGEERNPHVYQLLEGGEEETLRRLGLENNPOSYLYLVGQCAV	240
QY	257	SSINDKSPMKVYRKALSYIDFTEDEVEDLTSIVASVHLGNIHFAAEDSNAOVTTEENOL	316
DB	241	SSINDKSPMKVYRKALSYIDFTEDEVEDLTSIVASVHLGNIHFAAEDSNAOVTTEENOL	300
QY	317	KLTFTLLVEGEGTLREALTHKRIIAKGELLSPNLDAQAAVARDALAKAVYSRTFTWLVR	376
DB	301	KLTFTLLVEGEGTLREALTHKRIIAKGELLSPNLDAQAAVARDALAKAVYSRTFTWLVR	360
QY	377	KINRSLSKDAESPSMRSTVYGLLDYIGFVFOHNSFEOFCINCEKXQOLFETLTJK	436
DB	361	KINRSLSKDAESPSMRSTVYGLLDYIGFVFOHNSFEOFCINCEKXQOLFETLTJK	420

QY 437 SEOEYEAEGIAEPVOYFNKKIICDVEEKFGLIISIDDECLRGEATDLTLEKLED 496
 DB 421 SEOEYEAEGIAEPVOYFNKKIICDVEEKFGLIISIDDECLRGEATDLTLEKLED 480
 QY 497 TVKHPHFLTKLADOKTRKSLDRGEFRLHYAGEVTVSYTGFLDKNNDLFFNLKETMC 556
 DB 481 TIKHHPHFLTKLADOKTRKSLDRGEFRLHYAGEVTVSYTGFLDKNNDLFFNLKETMC 540
 QY 557 SSNMPMAOCPSKSELSDKKRPETVATOFKMSLLOLVEILRSKPEATYIRCIKPNDAKOPG 616
 DB 541 SSTPMMAOCPSKSELSDKKRPETVATOFKMSLLOLVEILRSKPEATYIRCIKPNDAKOPG 600
 QY 617 RFDEVLIRHGVKYLGLMENLVRRAAGFAYRRKRYEALQRYKSLCPEWPMAGRPDQVA 676
 DB 601 RFDEVLIRHGVKYLGLMENLVRRAAGFAYRRKRYEALQRYKSLCPEWPMAGRPDQVA 660
 QY 677 VLVRLHGYKPEEYMGKRTKIFIRPKTLFATEDSLVRSOSLAKTIOAARGFHMROKFL 736
 DB 661 VLVRLHGYKPEEYMGKRTKIFIRPKTLFATEDSLVRSOSLAKTIOAARGFHMROKFL 720
 QY 737 RVKSAICIOGMMWGTIGRRKAARKWAQAOTIRRLINGFLIRHSPCPENAFELDHVRAS 796
 DB 721 RVKSAICIOGMMWGTIGRRKAARKWAQAOTIRRLINGFLIRHSPCPENAFELDHVRAS 780
 QY 797 FLNLRROLPRNVLDTSWPTPPALREASSELLRELCKMNMWYKCRSISPEMKOOLQOKA 856
 DB 781 FLNLRROLPRNVLDTSWPTPPALREASSELLRELCKMNMWYKCRSISPEMKOOLQOKA 840
 QY 857 VASELFGKKNYPOSVRLFTSTRLGTEESIPRYLOSISLSEPIQYAVPVVYKRYKGYK 916
 DB 841 VASELFGKKNYPOSVRLFTSTRLGTEESIPRYLOSISLSEPIQYAVPVVYKRYKGYK 900
 QY 917 RPRQLLTPSAVVIYEDAKVKORIDYANLTGISVSSLSDFVLVHQBEDKKOGDVLQ 976
 DB 901 RSRQLLTPSAVVIYEDAKVKORIDYANLTGISVSSLSDFVLVHQBEDKKOGDVLQ 960
 QY 977 SDHVETLTKTALSADRVNNININGSTTFAGGPGRDIIDFTSGSELLITKAKNGHLAV 1036
 DB 961 SDHVETLTKTALSADRVNNININGSTTFAGGPGRDIIDFTSGSELLITKAKNGHLAV 1020
 QY 1037 VAPRLNSR 1044
 DB 1021 VAPRLNSR 1028
 RESULT 2
 A:59253
 C:Species: Homo sapiens (man)
 C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000
 C:Accession: A59253
 R:Clouet, F.; Amaraoui, A.E.; Blanchard, S.; Lenolr, M.; Ripoll, C.; Vago, P.; Hamel, C.;
 Genomics 40, 332-341, 1997
 A:Title: Cloning of the genes encoding two murine and human cochlear unconventional type
 A:Reference number: A59253; MIMD:97237053; PMID:9119401
 A:Accession: A59253
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 11028 <CRO>
 A:Cross-references: GB:X98507; NID:g1926310; PIDN:CA67131.1; PID:g1926311
 A:Experimental source: dev stage adult; tissue type kidney
 C:Genetics:
 A:Gene: myo-1b
 A:Map position: 17p3.2-p13.3
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 F:14-683/Domain: myosin motor domain homology <MMO>
 Query Match 95.3%; Score 5143; DB 2; Length 1028;
 Best Local Similarity 96.2%; Pred. No. 0;
 Matches 989; Conservative 23; Mismatches 16; Indels 0; Gaps 0;
 QY 17 MESALTRARVDVGVODVLENTSEAFIENLRPRRENLYTYIGPLVSVNPRDLOI 76
 I:|||||
 I:|||||

DB 1 MSALTRARVDVGVODVLENTSEAFIENLRPRRENLYTYIGPLVSVNPRDLOI 60
 QY 77 YSRQHMERYRGVSFEYEPHFLFAVDVYRALRTERDQAVMISSESGAGTEAKRLLQ 136
 DB 61 YSRQHMERYRGVSFEYEPHFLFAVDVYRALRTERDQAVMISSESGAGTEAKRLLQ 120
 QY 137 FYAECPAPRGGAVDRLQSNPYLFAFGNKTIRNDSSPFGYKMDVQDFKAPYVG 196
 DB 121 FYAECPAPRGGAVDRLQSNPYLFAFGNKTIRNDSSPFGYKMDVQDFKAPYVG 180
 QY 197 HILSYLLEKSRVYHONHGERNFHVFYOLLEGEETLRLGLERPNOSTLYLVKQCAKY 256
 DB 181 HILSYLLEKSRVYHONHGERNFHVFYOLLEGEETLRLGLERPNOSTLYLVKQCAKY 240
 QY 257 SSINDKSDKWKAKLSVDTFTDEVEDLSTYASVLHGNHFADESNQAVTTENOL 316
 DB 241 SSINDKSDKWKAKLSVDTFTDEVEDLSTYASVLHGNHFADESNQAVTTENOL 300
 QY 317 KYTLRLGVGEGTLREALTRHKIIAKGEBLSPNLQAAVARDALAKAVSRFTWLV 376
 DB 301 KYTLRLGVGEGTLREALTRHKIIAKGEBLSPNLQAAVARDALAKAVSRFTWLV 360
 QY 377 KINRSLASKDAESPMSRTVVLGLDIYGEVYQHNSTFOFCINTYCNKLOLFTLETLK 436
 DB 361 KINRSLASKDAESPMSRTVVLGLDIYGEVYQHNSTFOFCINTYCNKLOLFTLETLK 420
 QY 437 SEOEYEAEGIAEPVOYFNKKIICDVEEKFGLIISIDDECLRGEATDLTLEKLED 496
 DB 421 SEOEYEAEGIAEPVOYFNKKIICDVEEKFGLIISIDDECLRGEATDLTLEKLED 480
 QY 497 TVKHPHFLTKLADOKTRKSLDRGEFRLHYAGEVTVSYTGFLDKNNDLFFNLKETMC 556
 DB 481 TVKHPHFLTKLADOKTRKSLDRGEFRLHYAGEVTVSYTGFLDKNNDLFFNLKETMC 540
 QY 557 SSNMPMAOCPSKSELSDKKRPETVATOFKMSLLOLVEILRSKPEATYIRCIKPNDAKOPG 616
 DB 541 SSNMPMAOCPSKSELSDKKRPETVATOFKMSLLOLVEILRSKPEATYIRCIKPNDAKOPG 600
 QY 617 RFDEVLIRHGVKYLGLMENLVRRAAGFAYRRKRYEALQRYKSLCPEWPMAGRPDQVA 676
 DB 601 RFDEVLIRHGVKYLGLMENLVRRAAGFAYRRKRYEALQRYKSLCPEWPMAGRPDQVA 660
 QY 677 VLVRLHGYKPEEYMGKRTKIFIRPKTLFATEDSLVRSOSLAKTIOAARGFHMROKFL 736
 DB 661 VLVRLHGYKPEEYMGKRTKIFIRPKTLFATEDSLVRSOSLAKTIOAARGFHMROKFL 720
 QY 737 RVKSAICIOGMMWGTIGRRKAARKWAQAOTIRRLINGFLIRHSPCPENAFELDHVRAS 796
 DB 721 RVKSAICIOGMMWGTIGRRKAARKWAQAOTIRRLINGFLIRHSPCPENAFELDHVRAS 780
 QY 797 FLNLRROLPRNVLDTSWPTPPALREASSELLRELCKMNMWYKCRSISPEMKOOLQOKA 856
 DB 781 FLNLRROLPRNVLDTSWPTPPALREASSELLRELCKMNMWYKCRSISPEMKOOLQOKA 840
 QY 857 VASELFGKKNYPOSVRLFTSTRLGTEESIPRYLOSISLSEPIQYAVPVVYKRYKGYK 916
 DB 841 VASELFGKKNYPOSVRLFTSTRLGTEESIPRYLOSISLSEPIQYAVPVVYKRYKGYK 900
 QY 917 RPRQLLTPSAVVIYEDAKVKORIDYANLTGISVSSLSDFVLVHQBEDKKOGDVLQ 976
 DB 901 RSRQLLTPSAVVIYEDAKVKORIDYANLTGISVSSLSDFVLVHQBEDKKOGDVLQ 960
 QY 977 SDHVETLTKTALSADRVNNININGSTTFAGGPGRDIIDFTSGSELLITKAKNGHLAV 1036
 DB 961 SDHVETLTKTALSADRVNNININGSTTFAGGPGRDIIDFTSGSELLITKAKNGHLAV 1020
 QY 1037 VAPRLNSR 1044
 DB 1021 VAPRLNSR 1028
 RESULT 3
 S41749

myosin heavy chain I beta - bovine
N: Alternate names: myosin IB heavy chain
C: Species: Bos primigenius taurus (cattle)
C: Date: 20-May-1994 #sequence_revision 02-Jun-1995 #text_change 02-Feb-2001
C: Accession: S41749; A55744; S33497
R: Zihl, T.; Ikeda, M.
FEBS Lett. 339, 31-36, 1994

A: Title: A novel myosin I from bovine adrenal gland.
A: Reference number: S41749; MUID:94148088; PMID:8313976
A: Accession: S41749

A: Molecule type: mRNA

A: Residues: 1-1028 <ZHU>

A: Cross-references: GB:003420; MID:9436936; PIDN:AAA17565.1; PID:9436937

A: Experimental source: adrenal gland
R: Reizes, O.; Baryliko, B.; Li, C.; Suedhof, T.C.; Albanesi, J.P.
Proc. Natl. Acad. Sci. U.S.A. 91, 6349-6353, 1994

A: Title: Domain structure of a mammalian myosin Ibeta.

A: Reference number: A55744; MUID:94294379; PMID:8022785

A: Accession: A55744

A: Status: nucleic acid sequence not shown

A: Molecule type: mRNA

A: Residues: 1-6, 'C', 8-96, 'R', 98-319, 'N', 321-362, 'K', 364-385, 'D', 387-543, 'I', 545-549, 'F',
A: Cross-references: GB:22852; MID:9397402; PIDN:CAA80476.1; PID:9397403

A: Experimental source: brain
R: Reizes, O.; Li, C.; Suedhof, T.C.; Albanesi, J.P.
Submitted to the EMBL Data Library, May 1993

A: Description: Domain structure of a mammalian myosin I (bovine myosin IB).

A: Reference number: S33497

A: Accession: S33497

A: Molecule type: mRNA

A: Residues: 1-6, 'C', 8-96, 'R', 98-321, 'PST', 223-281, 'N', 283-296, 'GTSST', 303, 'PGPWNR', 307
A: Cross-references: GB:22852; MID:9397402; PIDN:CAA80476.1; PID:9397403

A: Experimental source: brain

C: Superfamily: brush border myosin heavy chain I; myosin motor domain homology
C: Keywords: actin binding; ATP; calmodulin binding; nucleotide binding; P-loop
F: 105-112/Region: nucleotide-binding motif A (P-loop)
F: 576-595/Region: actin binding #status predicted
F: 699-743/Region: calmodulin binding #status predicted
F: 111/Binding site: ATP (Lys) #status predicted

Query Match 95.0%; Score 5128; DB 2; Length 1028;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 983; Conservative 25; Mismatches 20; Indels 0; Gaps 0;

```

QY 17 MESALTARDRVGVDFVLENTSEAFIENLRFRFRFENLITTYIGPVLSVNPYRDLQ 76
DB 1 MESALTARDRVGVDFVLENTSEAFIENLRFRFRFENLITTYIGPVLSVNPYRDLQ 60
QY 77 YSRQHEBYRGVSFEVPPH.FAVADIVYRALRTERDDQVMSGESGAGKTEATRKLQ 136
DB 61 YSRQHEBYRGVSFEVPPH.FAVADIVYRALRTERDDQVMSGESGAGKTEATRKLQ 120
QY 137 FYAETCPAPERGAVERDRLQSNPVLEAFGNAKTLRDNDSRGKVDVDFDKGAPVG 196
DB 121 FYAETCPAPERGAVERDRLQSNPVLEAFGNAKTLRDNDSRGKVDVDFDKGAPVG 180
QY 197 HILSTLEKSRVHQNNGERNHFVYQLLEGGEETLRGLERPNQSYLYLVKGCACAY 256
DB 181 HILSTLEKSRVHQNNGERNHFVYQLLEGGEETLRGLERPNQSYLYLVKGCACAY 240
QY 257 SSINDKSDMKVYRKALSVIDETDEVEDLLSTYASVHLGNHFAADEDSNAQVTTENQ 316
DB 241 SSINDKSDMKVYRKALSVIDETDEVEDLLSTYASVHLGNHFAADEDSNAQVTTENQ 300
QY 317 KYTLRLGVETTLREALTHRKIIAKGELLSPLEQAAVYADALAKAVYSTRFTWLR 376
DB 301 KYTLRLGVETTLREALTHRKIIAKGELLSPLEQAAVYADALAKAVYSTRFTWLR 360
QY 377 KINSLSAKDAESPSMSTVVLGLDIYGFVFOHNSFEQFCINCNKRIQDLFTLTLK 436
DB 361 KINSLSAKDAESPSMSTVVLGLDIYGFVFOHNSFEQFCINCNKRIQDLFTLTLK 420

```

```

QY 437 SEQEEYEAEGIAPEVYFNNKIIICDLVEKEFGIISIDEBCLRGATDLTLEKLED 496
DB 421 SEQEEYEAEGIAPEVYFNNKIIICDLVEKEFGIISIDEBCLRGATDLTLEKLED 480
QY 497 TVKPHHFLTKLADOKTRKSLDRGEFLLHAYGEVYTVSGFLDKNNDLFRNLKEMC 556
DB 481 TVKPHHFLTKLADOKTRKSLDRGEFLLHAYGEVYTVSGFLDKNNDLFRNLKEMC 540
QY 557 SSMNIMAOCEFDKSESDKKRPETVATQFKMSLIDVETLRSEKREAYIRCIKPNAPG 616
DB 541 SSMNIMAOCEFDKSESDKKRPETVATQFKMSLIDVETLRSEKREAYIRCIKPNAPG 600
QY 617 RDEVLIRHOKVYKLGIMENLVRRAAGFAVRRKYEAFLOKYSICPETVPMAGRPQDVA 676
DB 601 RDEVLIRHOKVYKLGIMENLVRRAAGFAVRRKYEAFLOKYSICPETVPMAGRPQDVA 660
QY 677 VLVRLHGYKPEEKMGRTKIFTRPKTYLFTEDSLFVRQSLATKIQAMNGFHMROKFL 736
DB 661 VLVRLHGYKPEEKMGRTKIFTRPKTYLFTEDSLFVRQSLATKIQAMNGFHMROKFL 720
QY 737 RYKRSALCIQSWMRGTLGRKAKRMAAQTIRLRIRGFIIRHSPRCENAFIDHVRAS 796
DB 721 RYKRSALCIQSWMRGTLGRKAKRMAAQTIRLRIRGFIIRHSPRCENAFIDHVRAS 780
QY 797 FILNLRLQPLRVNLTSMPTPPALREASSELRLCKMNMWYKCRSISPEMKOOLQKA 856
DB 781 FILNLRLQPLRVNLTSMPTPPALREASSELRLCKMNMWYKCRSISPEMKOOLQKA 840
QY 857 VASEIFKGGKNDVPOSVPLFISTRGTGTEESPRVLOSISEPTIOYAVPVYKRYKYP 916
DB 841 VASEIFKGGKNDVPOSVPLFISTRGTGTEESPRVLOSISEPTIOYAVPVYKRYKYP 900
QY 917 RPROLLPSSAVYVEDAKVKORIDYANLTVISVSSLSDSLFLVHVRQEDKQKGVYLO 976
DB 901 RPROLLPSSAVYVEDAKVKORIDYANLTVISVSSLSDSLFLVHVRQEDKQKGVYLO 960
QY 977 SDHVIETLTKTSLADRVNNININGSTTFAGCGRCGCIIDFTSGSELLTKAKNGHLAV 1036
DB 961 SDHVIETLTKTSLADRVNNININGSTTFAGCGRCGCIIDFTSGSELLTKAKNGHLAV 1020
QY 1037 VAPRLNSR 1044
DB 1021 VAPRLNSR 1028

```

RESULT 4
151173
myosin I beta - bullfrog
C: Species: Rana catesbeiana (bullfrog)
C: Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Feb-2001
C: Accession: I51173
R: Metcalf, A.B.; Chelliah, Y.; Hudepeth, A.J.
Proc. Natl. Acad. Sci. U.S.A. 91, 11821-11825, 1994
A: Title: Molecular cloning of a myosin I beta isoform that may mediate adaptation by
A: Reference number: I51173; MUID:95083594; PMID:7991542
A: Accession: I51173
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-1028 <MET>
A: Cross-references: EMBL:U14549; MID:9602137; PIDN:AAA57192.1; PID:9602138
C: Geneticks:
A: Gene: AMI beta
A: Introns: 639/2
C: Superfamily: brush border myosin heavy chain I; myosin motor domain homology
C: Keywords: nucleotide binding; P-loop
F: 14-683/Domain: myosin motor domain homology <MOT>
F: 105-112/Region: nucleotide-binding motif A (P-loop)

Query Match 81.0%; Score 4371; DB 2; Length 1028;
Best Local Similarity 79.2%; Pred. No. 7; 8e-272;
Matches 814; Conservative 112; Mismatches 102; Indels 0; Gaps 0;

```

17 MESALTRADRVGVQVDFVLENTSEAFIENLRRENNLITYYIGPVLVSNPYRDLQI 76
1 MESALTRADRVGVQVDFVLENTSEAFIENLRRENNLITYYIGPVLVSNPYRDLQI 60
77 YSRQHMERYGVGFYEPVPHLEAVADTVYRALTERRDOQVMSGSGAGKTEATRRILQ 136
61 YSRQHMERYGVGFYEPVPHLEAVADTVYRALTERRDOQVMSGSGAGKTEATRRILQ 120
137 FYAETCAPRGAGVARDRLQSNPVLFAFNATKLRLNDSSRFKTYMDVQDFKGAAPVG 196
121 YVATGCVSDQVEVYKRRILQSNPVLFAFNATKLRLNDSSRFKTYMDVQDFKGAAPVG 180
197 HILSYLLEKSRVYVQNHGNERHFVYQLLEGSEETLRRLGRLERNPOSTYLKKGCAVY 256
181 HILSYLLEKSRVYVQNHGNERHFVYQLLEGSEETLRRLGRLERNPOSTYLKKGCAVY 240
257 SSINDKSDMKVMRKALSVIDFTEDEVEDLSIYASVHLGNHFADEDSNAQVTTENOL 316
241 SSINDKSDMKVMRKALSVIDFTEDEVEDLSIYASVHLGNHFADEDSNAQVTTENOL 300
317 KYLRLLGVEGTTLRREALTRHKRIIAKGEELLSPLNLEQAAYARDALAKAVYSRTFTWLYR 376
301 KYLRLLGVEGTTLRREALTRHKRIIAKGEELLSPLNLEQAAYARDALAKAVYSRTFTWLYR 360
377 KINSLASKDAESPMSSTVYGLDLYGFEVFOHNSFEOPCINVCNEKLOQLFIETLTK 436
361 KINSLASKDAESPMSSTVYGLDLYGFEVFOHNSFEOPCINVCNEKLOQLFIETLTK 420
437 SEOEYEAEGIAMEPVQYFNNKIIDLVEEKFGIISILDECLRPGEATDLFLEKLED 496
421 SEOEYEAEGIAMEPVQYFNNKIIDLVEEKFGIISILDECLRPGEATDLFLEKLED 480
497 TVKPHPHLTHKLADQTRKSLDGEFRLHYAGEVTVYTGFLDKNNDLFLRLKETMC 556
481 TVKPHPHLTHKLADQTRKSLDGEFRLHYAGEVTVYTGFLDKNNDLFLRLKETMC 540
557 SSMPNIAOCDKSELSDKRRPETVAVQFKMSLQVLVEILRSKRPATIRICRPDAKOPG 616
541 DSGNFIHOCFNRSLDTRKRPETVAVQFKMSLQVLVEILRSKRPATIRICRPDAKOPG 600
617 RFDEVLIHQVYKYLGLMNLVRRAGFAVRRKYEAFLORYKSLCPEPMWAGRPDQVGA 676
601 RFDEVLIHQVYKYLGLMNLVRRAGFAVRRKYEAFLORYKSLCPEPMWAGRPDQVGA 660
677 VLVRLHGYKPEEYKMGRTKIFIRPPKTLFATEDSLEVRQSLATKIQANRGGFWRKFL 736
661 VLVRLHGYKPEEYKMGRTKIFIRPPKTLFATEDSLEVRQSLATKIQANRGGFWRKFL 720
737 RVKRSALCIQSWMGRTGLRRKRAKKAQAQITRLIRGFLIRHSRCPENAFELDHVYAS 796
721 HHKHSAAVEIQSWMGRTGLRRKRAKKAQAQITRLIRGFLIRHSRCPENAFELDHVYAS 780
797 FLLNLRBQPRNVLDTSMPTRPPALREASSELLRELCMMWVKYCRSTISPEMKOOLQOKA 856
781 FLLNLRBQPRNVLDTSMPTRPPALREASSELLRELCMMWVKYCRSTISPEMKOOLQOKA 840
857 VASEIFKGGKKNYPOSVPRFLINTRLGNDEINTKILQOLESQTLTYAVPVVYDKRGYKP 916
841 VASEIFKGGKKNYPOSVPRFLINTRLGNDEINTKILQOLESQTLTYAVPVVYDKRGYKP 900
917 RPROLLITPSAVVYVEDAKVORIDYANLTGISVSSLSPLVLFVHVEDSNOKODVYLO 976
901 RPROLLITPSAVVYVEDAKVORIDYANLTGISVSSLSPLVLFVHVEDSNOKODVYLO 960
977 SDHVITLTKTALSDRVNNININOSITFAGPGRDGIIDFTSGSELLITKAKNGHLAV 1036
961 SDHVITLTKTALSDRVNNININOSITFAGPGRDGIIDFTSGSELLITKAKNGHLAV 1020
1037 VAPRLNSR 1044
1021 VAPRLNSR 1028

```

```

RESULT 5
H75634
myosin-1c - mouse (fragment)
N:Alternate names: myosin-I beta
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 08-Sep-2000
C:Accession: H75634
R:Crossref: F.; Amraoui, A.E.; Blanchard, S.; Lenoir, M.; Ripoll, C.; Vago, P.; Hamel,
Genomics 40, 332-341, 1997
A:Title: Cloning of the genes encoding two murine and human cochlear unconventional t
A:Reference number: A59253; M01D:97237053; PMID:9119401
A:Accession: H75634
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-807 <CRO>
A:Cross-references: GB:X9638; NID:91924960; PIDN:CA67956.1; PID:91924961
A:Experimental source: strain BALB/c; tissue type cochlea; dev stage adult
C:Genetics:
A:Gene: MGI:Myo1c
A:Cross-references: MGI:106612
A:Map position: 11:44.1
C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
F;14-683/Domain: myosin motor domain homology <MMO>

```

```

Query Match 77.2% Score 4167; DB 2; Length 807;
Best Local Similarity 99.1% Pred. No. 6,3e-259;
Matches 800; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 17 MESALTRADRVGVQVDFVLENTSEAFIENLRRENNLITYYIGPVLVSNPYRDLQI 76
DB 1 MESALTRADRVGVQVDFVLENTSEAFIENLRRENNLITYYIGPVLVSNPYRDLQI 60

QY 77 YSRQHMERYGVGFYEPVPHLEAVADTVYRALTERRDOQVMSGSGAGKTEATRRILQ 136
DB 61 YSRQHMERYGVGFYEPVPHLEAVADTVYRALTERRDOQVMSGSGAGKTEATRRILQ 120

QY 137 FYAETCAPRGAGVARDRLQSNPVLFAFNATKLRLNDSSRFKTYMDVQDFKGAAPVG 196
DB 121 FYAETCAPRGAGVARDRLQSNPVLFAFNATKLRLNDSSRFKTYMDVQDFKGAAPVG 180

QY 197 HILSYLLEKSRVYVQNHGNERHFVYQLLEGSEETLRRLGRLERNPOSTYLKKGCAVY 256
DB 181 HILSYLLEKSRVYVQNHGNERHFVYQLLEGSEETLRRLGRLERNPOSTYLKKGCAVY 240

QY 257 SSINDKSDMKVMRKALSVIDFTEDEVEDLSIYASVHLGNHFADEDSNAQVTTENOL 316
DB 241 SSINDKSDMKVMRKALSVIDFTEDEVEDLSIYASVHLGNHFADEDSNAQVTTENOL 300

QY 317 KYLRLLGVEGTTLRREALTRHKRIIAKGEELLSPLNLEQAAYARDALAKAVYSRTFTWLYR 376
DB 301 KYLRLLGVEGTTLRREALTRHKRIIAKGEELLSPLNLEQAAYARDALAKAVYSRTFTWLYR 360

QY 377 KINSLASKDAESPMSSTVYGLDLYGFEVFOHNSFEOPCINVCNEKLOQLFIETLTK 436
DB 361 KINSLASKDAESPMSSTVYGLDLYGFEVFOHNSFEOPCINVCNEKLOQLFIETLTK 420

QY 437 SEOEYEAEGIAMEPVQYFNNKIIDLVEEKFGIISILDECLRPGEATDLFLEKLED 496
DB 421 SEOEYEAEGIAMEPVQYFNNKIIDLVEEKFGIISILDECLRPGEATDLFLEKLED 480

QY 497 TVKPHPHLTHKLADQTRKSLDGEFRLHYAGEVTVYTGFLDKNNDLFLRLKETMC 556
DB 481 TVKPHPHLTHKLADQTRKSLDGEFRLHYAGEVTVYTGFLDKNNDLFLRLKETMC 540

QY 557 SSMPNIAOCDKSELSDKRRPETVAVQFKMSLQVLVEILRSKRPATIRICRPDAKOPG 616
DB 541 SSMPNIAOCDKSELSDKRRPETVAVQFKMSLQVLVEILRSKRPATIRICRPDAKOPG 600

QY 617 RFDEVLIHQVYKYLGLMNLVRRAGFAVRRKYEAFLORYKSLCPEPMWAGRPDQVGA 676
DB 601 RFDEVLIHQVYKYLGLMNLVRRAGFAVRRKYEAFLORYKSLCPEPMWAGRPDQVGA 660

QY 677 VLVRLHGYKPEEYKMGRTKIFIRPPKTLFATEDSLEVRQSLATKIQANRGGFWRKFL 736
DB 661 VLVRLHGYKPEEYKMGRTKIFIRPPKTLFATEDSLEVRQSLATKIQANRGGFWRKFL 720

```

```

Db      661 VLVRLHLYGKPEEYKMGRTKIFRPPKTLFATEDSLERRSLATKIOAARGHNMOKEL 720
OY      737 RKRSATCICQSMWGSTGCRKAARKKAQAOTIRRLINGFTLRSPRCPENAFLLDHYRAS 796
Db      721 RKRSATCICQSMWGSTGCRKAARKKAQAOTIRRLINGFTLRSPRCPENAFLLDHYRAS 780
OY      797 FLNLRLROLPRNVLDTSMPTRPPALRE 823
Db      781 FLNLRLROLPRNVLDTSMPTRPPALRD 807

```

RESULT 6

```

muslin IB - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 02-Feb-2001
C:Accession: S45574
R:Strom-Morgan, N.; Skovronsky, D.M.; Artavanis-Tsakonas, S.; Mooseker, M.S.
J. Mol. Biol. 239, 347-356, 1994
A:Title: The molecular cloning and characterization of Drosophila melanogaster myosin-I
A:Reference number: S45573; MUID:94260541; PMID:8201616
A:Accession: S45574
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1026 <STR>
A:Cross-references: GB:007596; NID:9466257; PIDN:AA19591.1; PID:9466258
C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
C:Keywords: nucleotide binding; P-loop
F:14-681/Domain: myosin motor domain homology <MOT>
F:105-112/Region: nucleotide-binding motif A (P-loop)

```

```

Query Match      48.8%; Score 2632; DB 2; Length 1026;
Best Local Similarity 51.5%; Pred. No. 2, 1e-160;
Matches 533; Conservative 177; Mismatches 300; Indels 24; Gaps 8;

OY      17 MESALTDARVGVQDVLLLENFTSEAAFTENRRRRRENLITTYIGPVLVSNPYRDLQI 76
Db      1 METGLERDAGVQDVELLENYQSEAFIGNLKRFOEDLITTYIGVLISVNPYKQLPI 60

OY      77 YSHOMERYRGVSEYVPHLPFVADTVYRALTRRPOAVVISESGAGTETKRLQ 136
Db      61 YTDNHVAKYRNKHFYEMPHIFAVYDNMFRSLIEENKQCVLISESSSGKTEASKYVQ 120

OY      137 FYAETCPAREGAVARDLLQSNPVLEAFGNAKTLIRNDSSRFKGYMDVDFGAPYVG 196
Db      121 FIAACGNGQTVGKVDKLLKSNPVLEAFGNAKTLIRNDSSRFKGYMDVDFGAPYVG 180

OY      197 HILSYLLKSRVYVHONGERNFHVYOLLEGESEETLRLGLERNPOSYLVVGCQAKV 256
Db      181 NILNLTLEKSRVYVADMGGERNFHITVQLLAGADEALQELGLERLDTYSTLTGLNGTV 240

OY      257 SSINDKSDKVKRKLASVIDETEDVLDLSTVAVLHGINHFADEDSNAQVTTENOL 316
Db      241 TSINDADSFQVQALTVYDFEGKEQREIFGIVASIHGLGVG-TEVEGNAKANSNDLV 299

OY      317 KVLTRLLEGSTLEALTLHRTIKAKGELLSPNLLEDAVAARALAKAVYSRTTWLVR 376
Db      300 VTAARLLGVNASEALTLHRTIDARGVVTSPLNQLALAYARALAKAVYDRLFSWLQ 359

OY      377 KINRSKASDAESPMSRSTVGLGLDYGFEVFOFPCINCYNEKLOOJFETLTLLK 436
Db      360 RLNLSLQAKETRA---SRNNVGLIDYGFIEIFOKNSFEOPCINFCNEKLOOJFETLTLLK 416

OY      437 SPOEYEAAGTAMEVOYFNKKITCDLVEEKFGIISLDEECLRPGEATDLTFLEKLD 496
Db      417 SDQDYRRREGLEWYVEYFNKVICNLLEEKHGIIISLDEECLRPGEPTDKTFLEKLD 476

OY      497 TVKPRHFLTKLADQKTRKSLDGEFFLHYAGEVYVYSGFDKNDLFLRKLKTM 556
Db      477 KLAQHHYVCHKAPAHIKKIMLDERFLVHYAGEVYVYNGFDKNDLFLRDLKETS 536

OY      557 SSNMPIMACDFKSELSDKRPETVANOQFKMSLLQVLEILRSKEPAYIRCIKPNDAKQPG 616

```

```

Db      537 KAGNGIVASCPPEKELSLKRPETAITQFRASLNNLMDLCKCEPSYIRIKRNDQITAN 596
OY      617 RDEVLIRHOVKTGLGMENTLVRAGATYRRKKEAFLOKRSKSCPTWPM--AGRPDQ 674
Db      597 VFNDELVHOVKTGLGMENTLVRAGATYRRKKEAFLOKRSKSCPTWPMYKPGCGKAG 656
OY      675 VAVLRLHLYGKPEEYKMGRTKIFRPPKTLFATEDSLERRSLATKIOAARGHNMOKEL 734
Db      657 VOQLVKDGMDEEYRVGTEKFLIRMPRTLFOTEDYAOEKKEITAIIOHNMGLMQRK 716
OY      735 FLRKRSATCICQSMWGSTGCRKAARKKAQAOTIRRLINGFTLRSPRCPENAFLLDHYR 794
Db      717 YLTKRAOYITMOSYCRKRLAOQAARKREAAKIRAFIKGFTIRNDAPNGFDEETANK 776
OY      795 ASFLNLRLROLPRNVLDTSMPTRPPALREASELLRELCKNMWKKCRISPEMKOQLO 854
Db      777 RMWTLRLAKELPTKVLDSKSPHAPGCEESGLHLRLHLRLARITRLTLPOOKROFEL 836
OY      855 KAVASEIFGKKNDYPOSVPRLEISTRLGTEEI---SPVLOSIGSEPIQYAVPVYKYDR 911
Db      837 KVLAEKVFKKKNVYASSVTFWFOEDRIPREHIQVNDPVASFGEOLKYSFCTKDR 896
OY      912 KGKPRPROLLRPSANVYVEDAKVQ-----RIDYANLGISVSSLSDSLFLVAVQR 964
Db      897 HGKRSRDRFTLLSNKATLYVDGTYRQKHLPLDKIDF-----TLTNHNDLAWYRIPL 950
OY      965 EDNRKQGDVYLOSDHYETLTKTALSADRVNINI-NQSYTFAGPGDGIIDTSSGE 1023
Db      951 DLKKDGDLLIPLRIESTYIIDTVGASIVSYDRNSLHNHYVKGAGVIDIQTGA 1010
OY      1024 LLITTKAKNGHLAV 1037
Db      1011 PGVVRDK-GHLVIT 1023

```

RESULT 7

```

B45438
myosin I beta, Mwi beta - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 31-Jan-2000
C:Accession: B45438
R:Sherr, E.H.; Joyce, M.P.; Greene, L.A.
J. Cell Biol. 120, 1405-1416, 1993
A:Title: Mammalian myosin I alpha, I beta, and I gamma: new widely expressed genes of
A:Reference number: A45436; MUID:93194946; PMID:8445986
A:Accession: B45438
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-448 <SHE>
A>Note: sequence extracted from NCBI backbone (NCBI:131911)
C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
F:1-448/Domain: myosin motor domain homology (fragment) <MOT>

```

```

Query Match      41.7%; Score 2252; DB 2; Length 448;
Best Local Similarity 97.1%; Pred. No. 1, 3e-136;
Matches 435; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY      162 LEAFNGAKTLIRNDSSRFKGYMDVDFDKAPVGGHILSYLLKSRVYVHONGERNHVF 221
Db      1 LEAFNGAKTLIRNDSSRFKGYMDVDFDKAPVGGHILSYLLKSRVYVHONGERNHVF 60

OY      222 YOLLEGESEETLRLGLERNPOSYLVYKGGCAVSSINDSKDKVKRKLASVIDETED 281
Db      61 YOLLEGESEETLRLGLERNPOSYLVYKGGCAVSSINDSKDKVKRKLASVIDETED 120

OY      282 VEDLSIYASVHLGINHFADEDSNAQVTTENOLKYLTRLGEGTTLREALTHHKRTA 341
Db      121 VEDLSIYASVHLGINHFADEDSNAQVTTENOLKYLTRLGEGTTLREALTHHKRTA 180

OY      342 KGEELSPNLNLEQAAVARDLAKAVYSRTTWLVRKINRSKASDAESPMSRSTVGLGL 401
Db      181 KGEELSPNLNLEQAAVARDLAKAVYSATETTWLVRKINRSKASDAESPMSRSTVGLGL 240

```


C:Keywords: actin binding; hydrolase; intestine; nucleotide binding; p-loop
 F:11-681/Domain: myosin motor domain homology <MOM>
 F:101-108/Region: nucleotide-binding motif A (p-loop)
 F:572-593/Region: actin binding #status predicted
 F:723-1043/Domain: carboxyl-terminal <CTD>
 F:107/Binding site: ATP (Lys) #status predicted

Query Match 39.0%; Score 2106.5; DB 1; Length 1043;
 Best Local Similarity 44.0%; Pred. No. 1e-126;
 Matches 455; Conservative 185; Mismatches 330; Indels 65; Gaps 21;

```

27  VGVDDVLENTSEAFIENLRRLRRRLIYIGPVLYSVNPRDQIYSRQHMYR 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8  VGVDDVLENTSEAFIENLRRLRRRLIYIGPVLYSVNPRDQIYSRQHMYR 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87  GVSFEVPHLEAVADTVRALTERRDQAVMISGSGAGKTEATRLQFAETCPAPE 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67  DYTTELKPHITLANNAYSLDRDRDQILITGSGAGKTEASKLVSVAACGKGE 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147  RGAADRRLQSNPVLEAFGNKTLRNDSRGKMYDQFDEKAPVGGHLSYLEKS 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127  QVNSVKEQLQSNPVLEAFGNKTLRNDSRGKMYDQFDEKAPVGGHLSYLEKS 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207  RVYHONGERNHFVYOLLEGEETLRRLGLERNPOSYLYLVKGQCAVSSINDKSMK 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187  RVYHONGERNHFVYOLLEGEETLRRLGLERNPOSYLYLVKGQCAVSSINDKSMK 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267  VMKRALSVIDFTEDEVEDLSTIYAVLHGNHFADEDSN----AQVTENQKLYTRL 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
246  VLQSAMTVIFSDDEIRQVLEVAALVTKGNVELNEFQANVPASGIDGQVEISEL 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
323  LGEVGTTLREALTHRKIIAKGELLSPNLEQAAVARDALAKAVYSRTFTWLVRKINSL 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
306  VGLNSVELEALQSLRMTAKERKVVYTLNVIAQVARDALAKAVYSRTFTWLVRKINSL 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
383  ASKDAESPWSRSTTVGLDIYGEFVFOHNSFEQNCINCKLQOULTELTSEQBEY 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
366  KVGTGE----KKVVGVDIYGFELLEDNSEFOYVINCNEKIQOVELTEMTLKEDEY 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443  EAGIAMEPVQYFNKKIICDLVEEKFGIISILDECLRPGEATDLTELEKEDTVKHP 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421  KRGIQVYVVEYEDNGIINLIEHNGRILAMDECLRPVVSSTFLAKINOLFSSHS 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503  HFLTHLADQKTRK---SLDRGEFRLHAYGEVTVSYGFLDKNNDLFRMLKETGSSM 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481  HY-ESVYTONAQROQIDHSMGLSCFRCHYAGKVTYVNSFDIKNNDLFRLSQAMMKAR 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
560  NPIMACFPKSE--LSDKRPETVATQFMSLQVETLRSEKPAVYICIRPNDAKQGR 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
540  HFLRLSLFPEQDPKQASLKRPTAGQFSSVTTLAMKLYSKNPYICIRPNEHQGRH 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
618  FDEVILRHQVKTGLGEMENLRVRAQFAYRRKYEAFLQRYKSLCEPTWMAAGPQDQAV 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
600  FSEFLVYQAOYLGLLENVRVRAQYAVROAVGSEFLERYRLRSSTWPMRNGQDEYEK 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
678  LVRLHLYKDEEYKMKRTKIFIRFPTLPTATEDSLEVRQSALTIQAMRFPNHRKQFLR 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
660  VIGELSMSEELAFKTKITFIRSPKTLTYLEBQRRLQOQATLTIQTYRGRCRTHQY 719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
738  VRSALCIQSNMGRGLGRRKAKRKAQAOTIRRLRGLRHSR-----CPENAF 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
720  MKRSQIVISVMEGRNMOKKHYRKMAASALLIOAFRGWKARKNRYKFRSQAALISNFI 779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
789  FLDAHVRASFLNLRQQLPR-NVLDTSWPTP-----PALRASLDELCKKNWVYCR 842
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
780  YKSMQV-KFLGLKNDLPSPSTLDKWPSSADYKYFNANHLQRLFQO-----MK-CK 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
843  ----SISPEMKQOLQOKAVASEIFKGDKNYPQSVRLFTIRGTETSPRYVLSGSE 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
831  KERDQSLSPKQVLEERKICASELFGKKASYPQSVPLPHFDYIGLQD-RNPKQKLGGE 889
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
899  --PIQYAVPVYKYNRKGKRRPRLDLPSSAVV--VEDAKVQRIDYANILGISVSSLS 954
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB 890 EGPILMAETVYKVRNGNAKTSRILLTKGHVITTDKMNPKQATVPIPLNSLAGVTSRK 949
 QY 955 DSLFVLAHQREDN-KQKGDVYQSDHYIELTK---TALSADVN-----NINIQ 1001
 DB 950 DGLSHLSSEISSGSGGEFLVSEHVEITLTKICRATLDATQMLPVYTEEFVSFKK 1009

Query Match 38.8%; Score 2094.5; DB 2; Length 1107;
 Best Local Similarity 41.7%; Pred. No. 6.6e-126;
 Matches 463; Conservative 163; Mismatches 346; Indels 119; Gaps 19;

```

27  VGVDDVLENTSEAFIENLRRLRRRLIYIGPVLYSVNPRDQIYSRQHMYR 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
15  IGVGDVYVLEPL-NEFTFIDNKKRDNHEITTYIGSVISVNPYKSLPTISPEKVEDYR 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87  GVSFEVPHLEAVADTVRALTERRDQAVMISGSGAGKTEATRLQFAETCPAPE 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
74  NRVFELSPHIALSDEARSLRDQDKCIIITGESGAGKTEASKLVMSVVAACGGA 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147  RGAADRRLQSNPVLEAFGNKTLRNDSRGKMYDQFDEKAPVGGHLSYLEKS 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134  EVNQVKEQLQSNPVLEAFGNKTYRNDSSRFEGKMYDEDFGDPGIVSYLLEKS 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207  RVYHONGERNHFVYOLLEGEETLRRLGLERNPOSYLYLVKGQCAVSSINDKSMK 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
194  RVYHONGERNHFVYOLLEGEETLRRLGLERNPOSYLYLVKGQCAVSSINDKSMK 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267  VMKRALSVIDFTEDEVEDLSTIYAVLHGNHFADEDSN----AQVTENQKLYTRL 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253  TVRNAAQIVGSDPEASEVLEVAALVTKGNVELNEFQANVPASGIDGQVEISEL 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
323  LGEVGTTLREALTHRKIIAKGELLSPNLEQAAVARDALAKAVYSRTFTWLVRKINSL 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
313  TSIDQVLELRARSFPTVEAKQEVSTYLVAAQATYARALAKKLYSLFSLVNRINESI 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
383  ASKDAESPWSRSTTVGLDIYGEFVFOHNSFEQNCINCKLQOULTELTSEQBEY 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
373  KQOTVVRK-----VMGVLDIYGFELLEDNSEFOIINYCNKRIQFIETLKEDEY 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443  EAGIAMEPVQYFNKKIICDLVEEKFGIISILDECLRPGEATDLTELEKEDTVKHP 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428  IREDLEWHIDYFNNAIICDLIENNTGILAMDECLRPGEATDLTELEKEDTVKHP 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503  HFLTHLADQKTRKSLDRGEFRLHAYGEVTVSYGFLDKNNDLFRMLKETGSSM 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
488  HESRMSKCSRFNDLTPHSCFRIOHYAGKVLQYVGGFADKNNDLRYRLDSQAMMKAGH 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
561  PIMACFPKSELS--DKRREPETVATQFMSLQVETLRSEKPAVYICIRPNDAKQGR 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 548 ALIKLEPEGNPAKVNLRKPPYAGSQTAKASVAILMKVLQTKNPVYICIRPNKKAHIF 607
 Oy 619 DEVLIRHQVKYLGIMENLRVRAGFAVRKRYEAFLOQKSLCPETWPMAGRPDGYAVL 678
 Db 608 SESLVCHQIRVLGLLENVRBAGYAFROAPECELEKMKLCQOTWPMHMGPARSGVEVL 667
 Oy 679 VRHLGYPEEKMKRTKIFIRFPTLATEDESLEVRQSLATKIQAAMRPFHMQKFLAV 738
 Db 668 FNELEIPVEEYSFGRSKIFIRNPRTLQLEDRLQRDLATLTQIKYRGKCTHFLM 727
 Oy 739 KRSALCISQSWRGTLGRKRAKRAKMAOTIRRLIRGF-----ILRHSPPCPENA---- 787
 Db 728 KRSQVVIAMWYRRRAQOKRYQOIKSSALVIGSYIRGMKARKILBELKHOKRCKEPAATIA 787
 Oy 788 -----FPLDH 792
 Db 788 AYWHGTQARRELKRLKEBARRRHAAVIAWYMLGLKYRREKFEFRANAGKRIYEFTLQR 847
 Oy 793 VRASFLMLRROLPR-NVLDTSMPPLPPALREASSELBELCMKNMW-----KYCRSISPEW 848
 Db 848 IVQKYLEMKMKMSPSLPIDKNWSPRYLFLDSTH--KELKRIFHLMRCKKYRQOFTDQO 905
 Oy 849 KOOLQOKAVASEIFKGRKNDYPOSVPLRFLISTRLGTEEISP-----RVLOSIGSEPIQYAV 904
 Db 906 KLIVEKLEASLFLKDKKALYPSVQCPQAGAYL--EINKNPYKKIKLDAIEEKIIIAE 962
 Oy 905 PYVKTRDKGYRPRQLLTFSAVIYE--DAKYQORIDYANLIGISVSSLSDELVLHV 962
 Db 963 VYNNKINRANGSTSFILFTNNLLLAQOKSGQISEVPLVDYTKVSMSSQNDGFFAVHL 1022
 Oy 963 QR-EDNKQGDVYVLOSDVHVIETLTK--TALSADRVN-NINI-----NOGSITFA 1007
 Db 1023 KGSSEAKSGDPLFSDHLEMAFKLYRTLSQTKQKLINEISDEFLVOFRQDKVCYKFI 1082
 Oy 1008 GSGPGDGIIDFTSGSELITAKKNGHLAVVA 1038
 Db 1083 QGNOKNGSVP-----TCRKNRNLLEVA 1105

RESULT 11

A45438
 myosin I alpha chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Feb-2001
 C:Accession: A45438
 R:Sherr, E.H.; Joyce, M.P.; Greene, L.A.
 J. Cell Biol. 120, 1405-1416, 1993
 A:Title: Mammalian myosin I alpha, I beta, and I gamma: new widely expressed genes of the
 A:Reference number: A45438; M01D:93194946; PMID:8449986
 A:Accession: A45438
 A:Status: Preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-1079 <SHE>
 A:Cross-references: GB:L00923; NID:9200008; PIDN:AAA39800.1; PID:9200009
 A:Experimental source: brain
 A:Note: Sequence extracted from NCBI backbone (NCBI:129047, NCBI:129048)
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C:Keywords: nucleotide binding; P-loop
 F:18-689/Domain: myosin motor domain homology <MMOT>
 F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match 38.6%; Score 2081.5; DB 2; Length 1079;
 Best Local Similarity 42.7%; Pred. No. 4.3e-125;
 Matches 463; Conservative 181; Mismatches 347; Indels 93; Gaps 20;

Oy 27 VGVODFVLENTFSEAFITENLRRENLIYIYIPVSVNYPRIQIYRQHEERYR 86
 Db 15 IGVGDVLEPL-NEETFLDNLKKRFDHNEIYIYISVIVNPRYSLSLTIYSEKEDYR 73
 Oy 87 GVSFYEVPPLHFAVDYVRAALTEERRDQAVMISGSGAGTEATKRLIQFAETCPAPE 146
 Db 74 NRPYELSPHIFALSEAVRSLRDQDKDCILITGSGAGKTEASKLMSVAAVCGKA 133

Oy 147 RGVAVDRLLQSNPVLEAFGNKTLTNDNSSRFKGYMDVOFDKCAPVGGHILSTLEKS 206
 Db 134 EVNVOKEQLQSNPVLEAFGNKTYVNDNSSRFKGYMDVOFDKCAPVGGHILSTLEKS 193
 Oy 207 RVVHONHGERNHFVQLLEGGEEPLRLGLERPNQSLYLKYKQCAKVSINDKSPK 266
 Db 194 RYVQCPGRGERNHFVQLLEGGEEPLRLGLERPNQSLYLKYKQCAKVSINDKSPK 252
 Oy 267 VMKRLSVIDEFDEVEDLLSIYASVILHGINIFAAD-----EDSNAQVTENQKYL 320
 Db 253 TVRNAMQYGFEDHEAEVLEVAVALKGINIEFKRESVNGIDESKIDKIELNEKFPAS 312
 Oy 321 RLIGVETTLREALTRKTIKANGELISPLNLEQAAAYADALAKAVYSTFTWYLRKINR 380
 Db 313 RPASVK-VVLERAFSEFVEAKREKYSTTLNVAQAAYADALAKAVYSTFTWYLRKINR 371
 Oy 381 SLASKDAESPSWSTVYGLDLYGFEVPOHNSFEQFCYNGCKEQLQFIETLTKSEOE 440
 Db 372 SIRAQTKVRKK-----VMGLDLYGFEIEDNSFEDQFIINCKEQLQFIETLTKSEOE 426
 Oy 441 EYEAGIAPWVQYFNKRIICDLVEERFKGIIISIDECELRPEATDTLPLEKLEDYVP 500
 Db 427 EYIRDEIEMTHIDYFNNAIICDLINNTNGIILAMDECELRPEATDTLPLEKLEDYVP 486
 Oy 501 HPHFLTH--KLADQTKRSLDRGEPLRHAAGVTVSYGTLDKNNDLFRLNKTETMCS 558
 Db 487 HOHEESMKSCKRFLNDTLPHSCFRQHTAGVLLVQBSGVKNNDLYRDSQAMKKA 546
 Oy 559 NNPIMACCFKSELS--DKRREPVAQFMKSLQVLEILRSKPEAYIRICIRPNDAKOPG 616
 Db 547 DHSILKLPFEGNPAKAKNLRPPYAGSQTAKASVAILMKVLQTKNPVYICIRPNKKAH 606
 Oy 617 RFDEVLIRHQVKYLGIMENLRVRAGFAVRKRYEAFLOQKSLCPETWPMAGRPDGYAV 676
 Db 607 IFNESLVCHOIRYVLGLLENVRBAGYAFROAPECELEKMKLCQOTWPMHMGPARSGVE 666
 Oy 677 VLVNHLGYPEEKMGRTKIFIRFPTLATEDESLEVRQSLATKIQAAMRPFHMQKFLAV 736
 Db 667 VLFNELEIPVEEYSFGRSKIFIRNPRTLQLEDRLQRDLATLTQIKYRGKCTHFLM 726
 Oy 737 RVKRSALCISQSWRGTLGRKRAKRAKMAOTIRRLIRGF-----ILRHSPPCPENA-- 787
 Db 727 LKRSQVVIAMWYRRRAQOKRYQOIKSSALVIGSYIRGMKARKILBELKHOKRCKEPAAT 786
 Oy 788 -----FPLDH 792
 Db 787 IAAVWGLKYRREKFEFRANAGKRIYEFTLQRIVQKYLEMKMKMSPSLPIDKNWSPR 846
 Oy 819 PALREASELRELCLCMKNMW-----KYCRSISPEWKOLOKRAVASEIFKGRKNDYPOSVPR 875
 Db 847 YLFLDSTH--KELKRIFHLMRCKKYRQOFTDQOKELEASSELFPKDKKALYPSVQO 904
 Oy 876 LFISTRGTEEISP-----RVLOSIGSEPIQYAVVYVYDKRGYRPRQLLTSPSAVIY 931
 Db 905 PFQAGAYL--EINKNPYKKIKLDAIEEKIIIAEYVNNKINRANGSTSFILFTNNLLLA 961
 Oy 932 E--DAKYQORIDYANLIGISVSSLSDELVLHVOR-EDNKQGDVYVLOSDVHVIETLTK-- 986
 Db 962 DQKSQISEVPLVDYTKVSMSSQNDGFFAVHLKESGSAASKGDFLSSDHLIMATKLY 1021
 Oy 987 -TALSADRVN-NINI-----NOGSITFAGPGDGIIDFTSGSELITAKKNGHL 1034
 Db 1022 RTLSQTKQKLINEISDEFLVOFRQDKVCYKFIQGNOKNGSVP-----TCRKNRNL 1073
 Oy 1035 AVVA 1038
 Db 1074 LEVA 1077

RESULT 12

A45439
 myosin I heavy chain - rat
 C:Species: Rattus norvegicus (Norway rat)

C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Feb-2001
 C/Accession: A45439; S29984
 R/Ruppert, C.; Kroschewski, R.; Bahler, M.
 J. Cell Biol. 120, 1393-1403, 1993
 A>Title: Identification, characterization and cloning of myr 1, a mammalian myosin-I.
 A/Reference number: A45439; MUID:93194945; PMID:8449985
 A/Accession: A45439
 A>Status: Preliminary
 A:Molecule type: mRNA; protein
 A:Residues: 1-1136 <RUP>
 A/Cross-References: GB:X68199; MID:956732; PIDN:CAA48287.1; PID:956733
 A/Note: Sequence extracted from NCBI backbone (NCBIN:128956, NCBI:128961)
 R/Ruppert, C.
 Submitted to the EMBL Data Library, September 1992
 A/Reference number: S29984
 A/Accession: S29984
 A>Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1136 <R02>
 A/Cross-References: EMBL:X68199; MID:956732; PIDN:CAA48287.1; PID:956733
 C/Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C/Keywords: nucleotide binding; P-loop
 F:108-688/Domain: myosin motor domain homology <MOT>
 F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match 38.5%; Score 2080; DB 2; Length 1136;
 Best Local Similarity 40.6%; Pred. No. 5.9e-125;
 Matches 463; Conservative 183; Mismatches 346; Indels 148; Gaps 19;

27 VGVQDFVLENTSEAFTEINLRRENTLYTYIGVLSVNPYDLOIYSRQHEMYR 86
 15 IGVGVTLLEPL-NEETFDNLKRRFDHNEYTYIGSVISVNPYRSLPIYSPEKVEDYR 73
 87 GVSFEVPEPLFAVADTVRALTERRDOAVMISEGAGTEAKRRLQFAETCPAPE 146
 74 NNFTELSPHITALDEAKRSRLDQDCILITGESAGKTEKSLVSYAAVCGKA 133
 147 RGVAVRDLRLSNPVLAEFGNAKTLRNDSSRFKYMVDQFPGKAPYGHILSTLEKS 206
 134 EVNQVKEQLQSTPLAEFNAKTVRNNDSSRFKYMVDQFPGKAPYGHILSTLEKS 193
 207 RVVHNGEERNHYVYQLLEGEEETLRRLERPOSTLYLVKGQCAKVSINDKDMK 266
 134 RVVQKPEERNHYVYQLLEGEEETLRRLERPOSTLYLVKGQCAKVSINDKDMK 252
 267 VMKALSYIDFTEDEVEDLSTIVASVLIHGNHFAADDSN---AQTTEGNOLKYLRL 322
 253 TVRNMQIVGFSDEPAESVLEVAVALKGINIEFRPESRMAGLDESKIKDKNELKEICEL 312
 323 LGVEETLRREALTHRKITAKEEELSLPLNLEDAAYARDALAKAVYSRTFTWLVKRINSL 382
 313 TSIDQVLEAFSEFTVEAKQEKVSTLTNAOAYARDALAKAVYSRTFTWLVKRINSL 372
 363 ASKDAESPSWSTVYGLDITGFVPHNSFEQPCITCYCEKIQOFTIELTKSEDEY 442
 373 KAQTVKRRK-----VMGLDITGFIEFEDNSFEQPCITCYCEKIQOFTIELTKSEDEY 427
 443 EAEGIAWEPVQYFNKNIICDLVEEKFQGIILDEECRPEATDLPFLKLEDTVPHP 502
 428 IREDIEWTHIDYFNNAITICDLLENNTNITLAMDDECRPSTVDEFFLEKLNOCAT 487
 503 HFLTH--KLADQTRKSLDGEFRLLAYAGVTVSVTGFELDKNDLLEFRNLKETMCSMN 560
 488 HFESMSKCSRFPLNDTLPNSCFRIQHYAGKVLQVEGFVDKNDLTYRDLQAMWAGH 547
 561 PIMACQFSKSELS--DKRREPVAPOFMSLLOVELIRSEKPAYICIRNDKAKQGRF 618
 548 ALISLFPENPAVANKRPPTAGSOKASAYATLKNQTKNPNYIKTIRKNDKKAHIF 607
 619 DEVLIRHQVYLLGLMELNVRRAQFAVRRKYEAFLOKYSKICPETWPMAGRPDGVAVL 678
 608 SESIVCHQIRYVLGLLENVRRAQFAVRRKYEAFLOKYSKICPETWPMAGRPDGVAVL 667

679 VRLHGYPEEYKMGKTKIFIRFPPTLFATEDSELEVRSLOSLATKIOAAMRGFHMROKELRV 738
 668 FNELEIPEVEEYSFGSKIFIRNPRTLFQEDLRQORLEDLATLQIKYRGKCKTHFLM 727
 739 KRSATCLOSMMRGTLGRKAKKAKKMAQOTIRLRIGF-----ILKHSPPCPNNA--- 787
 728 KRSQVYIAWYRRYAQQKRYQOIKSSALVYSYIRGMKARKILRELKHOKCKEAAVTIA 787
 788 ----- 787
 788 AYHNGTQARKRRRLKQDARKKHAIVYMAFVLSKARRELKRLKEARRKHAIVATWY 847
 788 -----FLDHYRASEFLMLRQLPR-NVLDTSWPTPPALR 822
 848 WGLKVRREYKFFRANAGKKIYFTLQRYQVLYLEKKNKMPSLPIDKNWPSRPYFL 907
 823 EASELRELCKKNWV---KYCRISPEMKQLOQKAVASIFPGKNDNPSPYRPFIS 879
 908 DSTH--KELKIFHLMKCKTRDOFTDOQKLIYEKLEASLEFKDKALPSSVGOFPQ 965
 880 TRLGTEELSP---RVLSLQSEPIQYAVPVYKRYRGRYPRPQLLTPSAVVIYV--D 933
 966 AYL---EINKNPYKTKLKDAIEEKIITAEVYKINRANGKSTSRIFLLTNNLLADQKS 1022
 934 AKVQRIDYANLPISTVSSLSDSLFLVLAQR-EDNKQGDVYLOSHVIETLR--TAL 989
 1023 GQISEVPLVDVTVMSSONDGFVAHLKESSEMAKSGDFLSSDLIMATKLYRTTL 1082
 990 SADRVN-NINI-----NQSITFAGPGDGIIIDTSGSELLITKAKNGHLAVVA 1038
 1083 SQTKQKNIETSEDFLVOFQDKYCVAFIQQNGKNSVP-----TCKRKNRLLEVA 1134

RESULT 13

S32404
 myosin heavy chain I, brain - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
 C/Accession: S32404; S33973
 R/Koslovsky, J.S.; Qian, C.; Jiang, X.; Mercer, J.A.
 FEBS Lett. 320, 121-124, 1993
 A>Title: Molecular cloning of a mouse myosin I expressed in brain.
 A/Reference number: S32404; MUID:93209370; PMID:8458427
 A/Accession: S32404
 A:Molecule type: mRNA
 A:Residues: 1-1094 <KOS>
 A/Cross-References: EMBL:X69987; MID:9312270; PIDN:CAA49604.1; PID:9166471
 R/Mercer, J.A.
 Submitted to the EMBL Data Library, January 1993
 A/Reference number: S33973
 A/Accession: S33973
 A:Molecule type: mRNA
 A:Residues: 1-572, 'G', 574-762, 'G', 764-1094 <MER>
 A/Cross-References: EMBL:X69987
 C/Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C/Keywords: nucleotide binding; P-loop
 F:5-675/Domain: myosin motor domain homology <MOT>
 F:95-102/Region: nucleotide-binding motif A (P-loop)

Query Match 38.4%; Score 2075.5; DB 2; Length 1094;
 Best Local Similarity 41.6%; Pred. No. 1.1e-124;
 Matches 462; Conservative 183; Mismatches 347; Indels 119; Gaps 19;

27 VGVQDFVLENTSEAFTEINLRRENTLYTYIGVLSVNPYDLOIYSRQHEMYR 86
 2 IGVGVTLLEPL-NEETFDNLKRRFDHNEYTYIGSVISVNPYRSLPIYSPEKVEDYR 60
 87 GVSFEVPEPLFAVADTVRALTERRDOAVMISEGAGTEAKRRLQFAETCPAPE 146
 74 NNFTELSPHITALDEAKRSRLDQDCILITGESAGKTEKSLVSYAAVCGKA 120
 147 RGVAVRDLRLSNPVLAEFGNAKTLRNDSSRFKYMVDQFPGKAPYGHILSTLEKS 206

121 EYNQVKEQLQSNPVLAEAFNAKTYRNDNSSRFQYMDIEFDKDDPLGCVISNILEKS 180
 207 RYVHONHGERNFHFVYOLLEGEETLRRLGLERNPOSYLIVKQCAKSSINDKSPK 266
 181 RYVKQPRGRNFHFVYOLLEGEETLRRLGLERNPOSYLIVKQCAKSSINDKSPK 239
 267 VAKKALSVDFEDEDLLISVAVLHGNTHFAEDDSN---AQTTEMLKYLRL 322
 240 TVRNAMQYVGFLDHAEVLAFLKGLNLEFESRNVNLDSEKIKDKKELEICEL 299
 323 LGEVETTLREALTHRIKAGEELSLPLNEQAARADALAKAVSRFTWLVKRNLSL 382
 300 TSDIYVLERASFPRVEAKQKYSTTLNVAQAYARADALANLXSRLESMVKNINSI 359
 383 ASKDAESPMSRSTVGLLDIYGFVFOHNSFOFCINCNEKLQOLFIETLTKSEOEY 442
 360 KATQVRRK-----VAGVLDIYGFIEFEDNSFEQFLINCNEKLQOLFIETLTKSEOEY 414
 443 EAGETAMEPVQYFNKKITCDLYEKFGLIISLDEECURRGATLTFTLEKEDYVKKP 502
 415 IREDLEWTHIDYFNNAITCDLIENTNGILAMDECELRGTVDLETLEKNOVCATHQ 474
 503 HELTH--KLADQTRKSLDGEFRLHYAGEVYVYSGFLDKNNDLERNLKETSCKSN 560
 475 HESRMSKCSRLNDTTLPHSCFRIOHAGVLYOGEFVDKNNDLRYDLISQAMWKGH 534
 561 PMAOCFDSKSELS--DKRPETVAOQKSLQVLEILSKREPATIRCKPNDAAQGRF 618
 535 SLTKSLFEPGNPAKVLKRPPTSSQFKASVATLNMNITOKNPNITRICKPNDKRAHLE 594
 619 DELVIRHOYKYLGLMENVRRAGFAFRKRYEAFLOKYSCLPETYPMAGRPDQVAVL 678
 595 NESLVCHOIRYGLLENVRRAGFAFRAGYEPCLERYCMLOKQWPHMKGPARBSVEVL 654
 679 VRHLGKPEYKMGRTKIFIRPKTLFATEDSLEROSGLATKIQAMFGFHWKQFLV 738
 655 FNLLEIRVEHSEHSGRSKIFIRNPRITFLEDLKORLEDLATLQIKYGMKCRHFLIM 714
 739 KRAIGIOSMWRGTLGRKRAKRAKMAQIIRLRIRF-----LRHSPRDEMA---- 787
 715 KRSQVIAAMRYEEOQKRIQOIKSALVIOSTIRGMKARKILRELKHOKKEAATTIA 774
 788 -----FELDH 792
 775 AYHGTQARRELKRLKEARANDAVIYAWMLGLKVRREYRKFRRANAGKKIYEPTLQR 834
 793 VRASFLLNLRQLR--NVDTSWPTPRPRLARESELLRELCKNMW---KXCRSISPEW 848
 835 IYQVYLLEKMKNMPSIDKMPSPRYLFLDSTH--KELKRIELHMRCKKYRDOPTDQO 892
 849 KOOLQOKAVASEIFKCKNDYPOSVPRFLFSTRLGTEESIP-----RVLOSLSSEPIQYAV 904
 893 KLIYEKLESELEFKKALYPSVQPOGATL---ELINKNPKYKKLDAIEKKIILAE 949
 905 PVAKYDRKGYKPPRROLLTPSAVIVE--DAVKQRIYANLTGIVSSLSDSLVLHV 962
 950 VVKIRANRANKSTSRIFLTLNNLLADQSGQIKSEVPLVDVTKVSSQNDGFEAVHL 1009
 963 QR-EDKKQGDVVLQDHYETITK---TALSADRN-NINI-----NQSITPA 1007
 1010 KESSEASAKGDFLSSDHLIEMATKLYRTLLSQTKKLNIEISDFLVOPRODKCVCKFI 1069
 1008 GGGPRDGIIDFTSGSELLITKANKGLAVVA 1038
 1070 QGNOKNSVP-----TCKRKNRNLLEVA 1092

RESULT 14
 A33620
 myosin heavy chain I, brush border - chicken (fragment)
 N:Contains: myosin ATPase (EC 3.6.4.1)
 C:Species: Gallus gallus (chicken)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002

C:Accession: A33620
 R:Garcia, A.; Coudrier, E.; Carboni, J.; Anderson, J.; Vandekerckhove, J.; Mooseker, M.
 J. Cell Biol. 109, 2895-2903, 1989
 A:Title: Partial deduced sequence of the 110-kD-calmodulin complex of the avian inces
 A:Reference number: A33620; MUID:90078325; PMID:2687288
 A:Accession: A33620
 A:Molecule type: mRNA
 A:Residues: 1-1000 <GAR>
 A:Cross-references: GB:X58479; NID:G63075; PIDN:CAA41388.1; PID:G63076
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C:Keywords: actin binding; hydrolase; intestine; nucleotide binding; P-loop
 F:1-639/Domain: myosin motor domain homology <MMO>
 F:59-66/Region: nucleotide-binding motif A (P-loop)
 F:530-551/Region: actin binding #status predicted
 F:681-1000/Domain: carboxyl-terminal <CTD>
 F:65/Binding site: ATP (uys) #status predicted

Query Match 35.3%; Score 1904; DB 1; Length 1000;
 Best Local Similarity 41.6%; Pred. No. 9.2e-114;
 Matches 415; Conservative 191; Mismatches 358; Indels 34; Gaps 16;

64 VLVSYPNPRDQIYTSRQMEKRYGVSYEPHFLFADVYVRLRERRDOAVMISGES 123
 2 VVISVNPKPPLIYTPKEVEYHNCNFEVAPATYALADATNSLRDRDQCLITGES 61
 124 GAGKTEATKRLQFVAETCPAPERGAVERDLQSNPVLAEAFNAKTYRNDNSSRFQYMD 183
 62 GAGKTEATKRLQFVAETCPAPERGAVERDLQSNPVLAEAFNAKTYRNDNSSRFQYMD 121
 184 DVQEDFKAPVGHILSTYLEKSRVYHONGERNFHFVYQOLLEGEETLRRLGLENNPQ 243
 122 DVEEPFKDPPDGGVYISNILEKSRIVRRKVEENRHFHLYQVLAGSQQLQOLLRDPCS 181
 244 SYIVLVKQCAKSSINDKSPKVRKALSYIDTFEDEDLLISVAVLHGNI---H 299
 182 HYGIY--NEKSVLQMGDAAAFRAMQDAMALIGRAPAEVRLLEVTAVVLKLGKVSLS 240
 300 FADEDSNAOYTTENQKLYLRLGVEGTTLREALTHRIKIAKEBELSLPLNEQAAYAR 359
 241 FOASGMEASIAEPRELOEISQILGIDPSTEOALCSRTVAVRDESVLTALYSQGYGR 300
 360 DALAAVYSRFTWLVKRNLSKASDAESPSWSTVYGLLDIYGFVFOHNSFEQFCI 419
 301 DALANNIYSRLFEDVKNRINTNSIOVKPG-----KQKVMGLDIYGFIEFODNGFEPFI 355
 420 NYCNEKLQOLFIETLTKSEOEYEAGETAMEPVQYFNKKITCDLYEKFGLIISLDEEC 479
 356 NYCNEKLQOLFIETLTKSEOEYEAGETAMEPVQYFNKKITCDLYEKFGLIISLDEEC 415
 480 LRGEATDITLEKLEDTYVKKPHFLHKLADQK--TRKSLDRGEFRLHYAGEVYVYSG 537
 416 LRGTAVNDTFTTKLNQIFASHKRYESKETLNAAHVVDVSLPLRCFRILHAGVTVYNT 475
 538 GLDKNNDLERNLKETSCKSNPMAOCFDSKSELS--DKRPETVAOQKSLQVLEILSK 595
 476 GETEKNNDLIRDLISQAMWAAARHTILBSLFEGBDQPRKSLKPTTOSQFASVATLMKN 535
 536 LSKSEPAYIRCIKPDNDKQPRGEVLLRHQVYKYLGLMENVRRAGFAFRKRYEAFLO 655
 536 LYSKNPNYIRCIKPDNDKQPRGEVLLRHQVYKYLGLMENVRRAGFAFRKRYEAFLO 595
 656 YKSLCPETWPMAGRPDQVAVVHNLGKPEYKMGRTKIFIRPKTLFATEDSLER 715
 536 YKSLCPETWPMAGRPDQVAVVHNLGKPEYKMGRTKIFIRPKTLFATEDSLER 715
 536 YKSLCPETWPMAGRPDQVAVVHNLGKPEYKMGRTKIFIRPKTLFATEDSLER 715
 716 QSLATKIOAAMRGFMRQFLKVRSAICIOSMWRGTLGRKRAKRAKMAQIIRLRIRF 775
 656 AELATLQKMFQMGCCCKRYQLMKRSQILLISAMFGHQRRNRYQMKRSVLLQAYARGW 715
 776 IIR-----HSRQPE--NAFFLDHVRASFLINLRQI--PRNVDTSWPTPR--PALKAS 825
 716 KTRRMVRRYFRSDACTRLSNFIYRBMVOYKIMGLQKNLPMAVILDRTPAPARYFLSDAN 775

QY 82 ELLRELCMMNMWKKCRSLSPKMKOOLQOKVAVASIEFKKKDNPQGVPRFLISTRG-T 884
Db 776 QELKSTFFYMKCKKREOLTPQORMLAKMLPRLFDKRALYQSLQOPFGEYIGLT 835
QY 885 EELSPPLVLOSLSGSEPIQYAVPVKKDKRGYKPRPRLTLITPRAVIVED--AKKQGITDY 942
Db 836 QNKYKGLQAVAKDKRYMAEVAQYKNRANGKTVPRLLTLTEHLVLDPRKAQYKMWLSL 895
QY 943 ANLTGISVSSLSDSLEFLVHVOREDNK-QKGDVVLQSDHVIEITLTK---TALSAD-RVNNI 997
Db 896 CDIOGASVRSFSGLLALHLKETSTAGCKGDDLTPSLILIELVRLHQTLLMDATAQALPL 955
QY 998 NI-NOGSITFAGGPGMDGIIDFT-SGSELLTKKANGH 1033
Db 956 SIADQFSTRPKGDVAALVTAVESAKGQDPVCKKRGSH 993

```

RESULT 15
A:53933
myosin I myr 4 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 02-Feb-2001
C:Accession: A53933
R:Baehlor, M.; Kroschewski, R.; Stoeffler, H.E.; Behrmann, T.
J. Cell Biol. 126, 375-389, 1994
A:Title: Rat myr 4 defines a novel subclass of myosin I: Identification, distribution,
A:Reference number: A53933; MUID:94308268; PMID:8034741
A:Accession: A53933
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-1006 <BAE>
A:Cross-references: GB:X71997; NID:q516395; PID:q516396
C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
C:Keywords: nucleotide binding, P-loop
F:12-682/Domain: myosin motor domain homology <MMOT>
F:102-109/Region: nucleotide-binding motif A (P-loop)

```

Query Match	34.5%;	Score 1861;	DB 2;	Length 1006;
Best Local Similarity	41.7%;	Pred. No. 5.3e-11;		
Matches 426;	Conservative 158;	Mismatches 301;	Indels 136;	Gaps 25;

[illegible][illegible]

Search completed: July 7, 2003, 14:24:54
Job time : 59.1849 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:44 ; Search time 29.5472 Seconds

(without alignments)
1465.497 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398
Sequence: 1 MRYRASALSGDGVRTMESA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5320	98.6	1028	1 MYIC_MOUSE	Q9WCI7 mus musculu
2	5143	95.3	1028	1 MYIC_HUMAN	Q00159 homo sapien
3	2626	48.6	1026	1 MYIB_DROME	Q23979 drosophila
4	2106.5	39.0	1043	1 MYHL_BOVIN	P10568 bos taurus
5	2090.5	38.7	1107	1 MYLA_MOUSE	P46735 mus musculu
6	2080	38.5	1136	1 MYLA_MOUSE	Q05096 rattus norv
7	2064.5	38.2	1043	1 MYHL_HUMAN	Q9UDC5 homo sapien
8	2005.5	37.2	1045	1 MYHL_CHICK	P47807 gallus gall
9	1987	36.8	909	1 MYHL_MOUSE	Q88329 mus musculu
10	1932.5	35.8	842	1 MYHL_RAT	Q62774 rattus norv
11	1800.5	33.4	1011	1 MYLA_DROME	Q23978 drosophila
12	1739	32.2	1003	1 MYSE_MOUSE	Q03479 dictyostell
13	1726.5	32.0	1099	1 MYSE_MOUSE	P70248 mus musculu
14	1692.5	31.4	1111	1 MYSD_MOUSE	P34092 dictyostell
15	1601	29.7	1113	1 MYSD_DICDI	P22467 dictyostell
16	1588.5	29.4	994	1 MYSA_DICDI	P34109 dictyostell
17	1574.5	29.2	1147	1 MYSB_ACACA	P19706 acanthamoeb
18	1523.5	28.2	1168	1 MYSC_ACACA	Q04439 saccharomyc
19	1501	27.8	1219	1 YMZ9_YEAST	P42522 dictyostell
20	1453	26.9	1181	1 MYSC_YEAST	P36006 saccharomyc
21	1427.5	26.4	1273	1 MYSA_HUMAN	Q13402 homo sapien
22	1354	25.1	2215	1 MYSA_HUMAN	P79114 bos taurus
23	1325	24.5	2052	1 MYLO_BOVIN	P97479 mus musculu
24	1324.5	24.5	2215	1 MYLA_MOUSE	Q96424 mus musculu
25	1284.5	23.8	2058	1 MYLA_MOUSE	Q96424 mus musculu
26	1279	23.7	2116	1 MYLA_MOUSE	Q96424 mus musculu
27	1263	23.4	1742	1 MYSC_DICDI	Q96424 mus musculu
28	1259	23.3	3511	1 MYIS_HUMAN	Q96424 mus musculu
29	1243.5	23.0	3530	1 MYIS_MOUSE	Q96424 mus musculu
30	1233.5	22.9	1509	1 MYIS_HUMAN	Q96424 mus musculu
31	1225.5	22.7	1853	1 MYSA_MOUSE	Q96424 mus musculu
32	1224.5	22.7	1828	1 MYSA_MOUSE	Q96424 mus musculu
33	1224.5	22.7	2245	1 MYSD_DICDI	Q96424 mus musculu

34	1217.5	22.6	1938	1 MYA_AEOR	P24733 aequipecten
35	1217	22.5	1855	1 MYA_HUMAN	Q96411 homo sapien
36	1216.5	22.5	1846	1 MYB_RAT	P70569 rattus norv
37	1206	22.3	1829	1 MYSA_CHICK	Q02440 gallus gall
38	1206	22.3	1978	1 MYH_CHICK	P10587 gallus gall
39	1202	22.3	1935	1 MYH_PIG	P79293 sus scrofa
40	1198	22.2	1935	1 MYH_RAT	P02564 rattus norv
41	1192	22.1	1972	1 MYH_HUMAN	P35749 homo sapien
42	1192	22.1	2017	1 MYSN_DROME	Q99323 drosophila
43	1191	22.1	1939	1 MYH_HUMAN	P13533 homo sapien
44	1190	22.0	1972	1 MYH_RABIT	P35748 oryctolagus
45	1187.5	22.0	1976	1 MYA_BOVIN	Q27991 bos taurus

ALIGNMENTS

RESULT 1
MYIC_MOUSE STANDARD: PRT: 1028 AA.
ID MYIC_MOUSE
AC Q9WCI7; 008834; 008571; Q9QW54;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin Ic (Myosin I beta) (MIMD).
GN MYOIC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC STRAIN=BA/2J; TISSUE=Brain;
RX MEDLINE=97325785; PubMed=9182797;
RA Hamilton B.A., Smith D.J., Mueller K.L., Kerrebrock A.W.,
RA Bronson R.T., van Berkel V., Daly M.J., Kruglyak L., Reeve M.P.,
RA Nemhauser J.L., Hawkin T.L., Rubin E.M., Lander E.S.,
RT "The vibrator mutation causes neurodegeneration via reduced expression
RT of PTP alpha: positional complementation cloning and extragenic
RT suppression."
RT Neuron 18:711-722(1997).
RL [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawaji T., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boiffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kanliya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima M., Rodriguez I., Sakamoto N.,
RA Nordone P., Rong B., Rongwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokawa K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RT Nature 405:685-690(2001).
RL [3]
RP SEQUENCE OF 1-807 FROM N.A. (ISOFORM B).
RC STRAIN=BA/2J; TISSUE=Cochlea;
RX MEDLINE=97237053; PubMed=9119401;
RA Crozet F., Amroui A.E., Blanchard S., Lenoir M., Ripoll C., Vago P.,
RA Hanel C., Fitzmaes C., Levi-Acobas F., Depetris D., Mattei M.-G.,
RA Weil D., Pujol R., Petit C.,
RT "Cloning of the genes encoding two murine and human cochlear

RT unconventional type I myosins.";
 RL Genomics 40:332-341(1997).
 RN [4]
 RP SEQUENCE OF 146-593 FROM N.A.
 RX MEDLINE-93194946; PubMed-8449986;
 RA Sherr E.H., Joyce M.P., Greene L.A.;
 RT "Mammalian myosin I alpha, I beta, and I gamma: new widely expressed
 RL genes of the myosin I family."
 J. Cell Biol. 120:1405-1416(1993).
 CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
 CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
 CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
 CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
 CC SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IO DOMAINS.
 CC -----
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U96723; AAC53264.1; -
 CC EMBL: AK004743; BAB23524.1; -
 CC EMBL: U96726; AAC60758.1; -
 CC EMBL: X99638; CAAG7956.1; -
 CC HSSP: P08799; 1MND.
 CC MGD: MGI:106612; Myo1c.
 CC InterPro: IPR000048; IO_region.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF00063; myosin_head.
 CC Pfam: PF00612; IO_3.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC ProDom: PD000355; myosin_head; 1.
 CC SMART: SM00015; IO_2.
 CC SMART: SM00242; MYSC; 1.
 CC PROSITE: PS50096; IO; 2.
 CC Myosin: ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
 CC Multigene family; Alternative splicing.
 CC KW DOMAIN 1 683
 CC FT 699 722
 CC FT 723 751
 CC FT 103 112
 CC FT NF_BIND
 CC VARSPLIC 897 1028
 CC
 CC RR -> GG (IN REF. 3).
 CC SKO -> RKR (IN REF. 3).
 CC T -> A (IN REF. 4).
 CC C -> F (IN REF. 4).
 CC VM -> LL (IN REF. 4).
 CC R -> A (IN REF. 4).
 CC LAS -> VPA (IN REF. 4).
 CC Q -> R (IN REF. 4).
 CC VKP -> IKR (IN REF. 4).
 CC M -> T (IN REF. 4).
 CC S -> G (IN REF. 4).
 CC O -> R (IN REF. 3).
 CC R -> G (IN REF. 3).
 CC E -> D (IN REF. 3).
 CC
 CC 1028 AA; 118155 MW; 2CF5535D816884DF CRC64;
 SO SEQUENCE

Query Match 98.6%; Score 5320; DB 1; Length 1028;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1028; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	17	MESALTARDRYGVDFVLENTSEAFIENLRREFRENTLYTIGFVLSVNPYRDLQI	76
DB	1	MESALTARDRYGVDFVLENTSEAFIENLRREFRENTLYTIGFVLSVNPYRDLQI	60
QY	77	YSROHMERXRCVSFEVPPHLLFAVADTVYRALFTRERDQAVMISGESGAGTAYRKLLO	136
DB	61	YSROHMERXRCVSFEVPPHLLFAVADTVYRALFTRERDQAVMISGESGAGTAYRKLLO	120
QY	137	FYACTCAPENGAVERPRLLOSNPVLAFAFGAKTLRDNSSRGKYADVDFDKGAPVYG	196
DB	121	FYACTCAPENGAVERPRLLOSNPVLAFAFGAKTLRDNSSRGKYADVDFDKGAPVYG	180
QY	197	HILSYLEKSRVYVQNHGERNFHFYQLLEGEEETLRRLGLENNPOSYLYVKGQCAKY	256
DB	181	HILSYLEKSRVYVQNHGERNFHFYQLLEGEEETLRRLGLENNPOSYLYVKGQCAKY	240
QY	257	SSINDKSDMKVKKALSYIDTFEDEVEDLSIVASVYHLGNIHRAEDSNQAVTENNOL	316
DB	241	SSINDKSDMKVKKALSYIDTFEDEVEDLSIVASVYHLGNIHRAEDSNQAVTENNOL	300
QY	317	KYLRLRGVEGTTLRALTRHKITAKGELLSPNLBOAAYARDALAKAYSRFTWLV	376
DB	301	KYLRLRGVEGTTLRALTRHKITAKGELLSPNLBOAAYARDALAKAYSRFTWLV	360
QY	377	KINSLASKDAEPPSWRSTVGLDLYGFEVFQHNSEFQPCINCEKLOQFETLT	436
DB	361	KINSLASKDAEPPSWRSTVGLDLYGFEVFQHNSEFQPCINCEKLOQFETLT	420
QY	437	SEQEEYEAEGTAMEPVQYFNKKIICDLYEEKFKIISTIDECRPRGEATDPLFLEKED	496
DB	421	SEQEEYEAEGTAMEPVQYFNKKIICDLYEEKFKIISTIDECRPRGEATDPLFLEKED	480
QY	497	TYKPRPHFLTKLADOKTRKSLDRGEFRLHYAGEVTVSYTGFLDKNNDLFNNLKETWC	556
DB	481	TYKPRPHFLTKLADOKTRKSLDRGEFRLHYAGEVTVSYTGFLDKNNDLFNNLKETWC	540
QY	557	SSANPIMACPDKSLSKKRPETVAYQFKSLQVYLLRSKPAYIRCIKPPADAKOPG	616
DB	541	SSANPIMACPDKSLSKKRPETVAYQFKSLQVYLLRSKPAYIRCIKPPADAKOPG	600
QY	617	RFDEVLIHROYKYLGLIMENLRVRAGFAVRRKYEAFLORYKSLCPETWPMAGRPDQVA	676
DB	601	RFDEVLIHROYKYLGLIMENLRVRAGFAVRRKYEAFLORYKSLCPETWPMAGRPDQVA	660
QY	677	VIVRHLGYKPEEYKMGRTKIFIRPKTLFATEDSLVVRQSLATKIQAAWGFHWKQFL	736
DB	661	VIVRHLGYKPEEYKMGRTKIFIRPKTLFATEDSLVVRQSLATKIQAAWGFHWKQFL	720
QY	737	RVKRSALICQSWMRGTIGRRKAAKRWMAOTIRLIRGFILRHSPRCENAFFLDHYRAS	796
DB	721	RVKRSALICQSWMRGTIGRRKAAKRWMAOTIRLIRGFILRHSPRCENAFFLDHYRAS	780
QY	797	FLILNRROLPRNVLDTSWTPPALREASSELLRELCMKNMWYKCRSISPWKQOLOOKA	856
DB	781	FLILNRROLPRNVLDTSWTPPALREASSELLRELCMKNMWYKCRSISPWKQOLOOKA	840
QY	857	VASEIFPKGKNDYPOSVPRLLISTRLGTEBISPRVLOSLGSEPPQYANPVYKVRKGKXP	916
DB	841	VASEIFPKGKNDYPOSVPRLLISTRLGTEBISPRVLOSLGSEPPQYANPVYKVRKGKXP	900
QY	917	RRROLLTPPSAVIYEDAKVKORIDYANLTGTSVSLDSFLVHVQREDNKKOGDVYLO	976
DB	901	RRROLLTPPSAVIYEDAKVKORIDYANLTGTSVSLDSFLVHVQREDNKKOGDVYLO	960
QY	977	SDHVLETLTKTALSADRVNNININGSTIFAGPGRGDIIDFTSGSELLITKANKGLAV	1036
DB	961	SDHVLETLTKTALSADRVNNININGSTIFAGPGRGDIIDFTSGSELLITKANKGLAV	1020
QY	1037	VAPRLNSR 1044	
DB	1021	VAPRLNSR 1028	


```

RESULT 2
MYIC_HUMAN STANDARD; PRT; 1028 AA.
ID MYIC_HUMAN
AC 000159;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin IC (Myosin I beta) (MMI-beta) (MMIB).
GN MyoIC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97237053; PubMed=9119401;
RA Crozet F., Amroult A.E., Blanchard S., Lenoir M., Ripoll C., Vago P.,
RA Hamel C., Fitzames C., Levi-Acobas F., Depetris D., Mattei M.-G.,
RA Weil D., Pujol R., Petit C.;
RT "Cloning of the genes encoding two murine and human cochlear
RT unconventional type I myosins.";
RL Genomics 40:332-341(1997).
CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC
DR EMBL: X8507; CAA67131.1; -.
DR HSSP: P08799; 1MDN.
DR GeneW: HGNC:7596; MYO1B.
DR GeneW: HGNC:7597; MYO1C.
DR MIM: 606538; -.
DR
DR InterPro: IPR000048; IQ_region.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00612; IQ; 3.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 2.
DR SMART: SM00242; MISC; 1.
DR PROSITE: PS50096; IQ; 2.
KW Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
KW Multigene family.
FT DOMAIN 1 683 MYOSIN HEAD-LIKE.
FT DOMAIN 699 722
FT DOMAIN 723 751 IQ 1.
FT DOMAIN 751 772 IQ 2.
FT NP_BIND 105 112 ATP (POTENTIAL).
SQ SEQUENCE 1028 AA; 118037 MW; 0E9C3680527F85C6 CRC64;
Query Match 95.3%; Score 5143; DB 1; Length 1028;
Best Local Similarity 96.2%; Pred. No. 5.6e-319;
Matches 989; Conservative 23; Mismatches 16; Indels 0; Gaps 0;

```

```

OY 17 MESALTARDRGVDFVLENTSEAFIEMLRRFRFENLTYTYIGPVLSVNPYRDLQI 76
DB 1 MDSALTARDRGVDFVLENTSEAFIEMLRRFRFENLTYTYIGPVLSVNPYRDLQI 60
OY 77 YSRQHEMYRGVSFEVPPHFAVADYTRALTRERKQAVMISGESGAGTEATKRLQ 136
DB 61 YSRQHEMYRGVSFEVPPHFAVADYTRALTRERKQAVMISGESGAGTEATKRLQ 120
OY 137 FYAETCPADRGCAVRDLQSNPVLEAFGNKTLRDNSSRFGRKYMVQDFPKGAPYGG 196
DB 121 FYAETCPADRGCAVRDLQSNPVLEAFGNKTLRDNSSRFGRKYMVQDFPKGAPYGG 180
OY 197 HILSYLLEKSRVHONHGRNHFVYQLLEGESEETLRGLERNPQSYLYLVKQCAKV 256
DB 181 HILSYLLEKSRVHONHGRNHFVYQLLEGESEETLRGLERNPQSYLYLVKQCAKV 240
OY 257 SSINDSDMKVKKALSVIDFTEDEVEDLSTIVASVLIHGNHFAADSDNAQVYTEQL 316
DB 241 SSINDSDMKVKKALSVIDFTEDEVEDLSTIVASVLIHGNHFAADSDNAQVYTEQL 300
OY 317 KILTRLLVEGTTLEALTHKRIKAGBELLSPLNEQAAVARDALAAVYSRFTWLVYR 376
DB 301 KILTRLLVEGTTLEALTHKRIKAGBELLSPLNEQAAVARDALAAVYSRFTWLVYR 360
OY 377 KINRSIASDAEPMRSTVYGLLDYGFVFOHNSPEOCINCKEQLDFTLETLK 436
DB 361 KINRSIASDAEPMRSTVYGLLDYGFVFOHNSPEOCINCKEQLDFTLETLK 420
OY 437 SEQEEYEAEGIAEPEVQYFNKKIICDLVEEKFGLIISLDEECURPGATDFTLEKLE 496
DB 421 SEQEEYEAEGIAEPEVQYFNKKIICDLVEEKFGLIISLDEECURPGATDFTLEKLE 480
OY 497 TYKPHPHLTHKRLADOKTRKSLDGEFFLLHYAGVYTSYGFIDKNDLLFRULKETMC 556
DB 481 TYKPHPHLTHKRLADOKTRKSLDGEFFLLHYAGVYTSYGFIDKNDLLFRULKETMC 540
OY 557 SSMPDMAOCPDKSELSDKKRPETVATQFKKSLDLVLEILSKPEAYIRCIKPDANOPG 616
DB 541 SSMPDMAOCPDKSELSDKKRPETVATQFKKSLDLVLEILSKPEAYIRCIKPDANOPG 600
OY 617 RFEDEVLIHQVYKYLGLMENLVRNAGFAVRRKYAFLOKRYKSLCPETWMAAGRPQDVA 676
DB 601 RFEDEVLIHQVYKYLGLMENLVRNAGFAVRRKYAFLOKRYKSLCPETWMAAGRPQDVA 660
OY 677 VLVRHLGKPEEYKMGRTKIFRPPKTLFAFEDSLVYRROSLAKTIOAMNGEHWKQFL 736
DB 661 VLVRHLGKPEEYKMGRTKIFRPPKTLFAFEDSLVYRROSLAKTIOAMNGEHWKQFL 720
OY 737 RVKRSALICQSMWNGTIGRRRAAKRKMAAQIRLLINGFILRHSPRCPENAFELDHYRAS 796
DB 721 RVKRSALICQSMWNGTIGRRRAAKRKMAAQIRLLINGFILRHSPRCPENAFELDHYRAS 780
OY 797 FLNLRLQPLRNVDLTSMPRPALREASSELRLCLCKNMWKCISPEMKOOLQOKA 856
DB 781 FLNLRLQPLRNVDLTSMPRPALREASSELRLCLCKNMWKCISPEMKOOLQOKA 840
OY 857 VASEIFGKNDYPOSVPRLFISTRLGTEISPVLSLSEPIQYAVPVVKKYDRKGYR 916
DB 841 VASEIFGKNDYPOSVPRLFISTRLGTEISPVLSLSEPIQYAVPVVKKYDRKGYR 900
OY 917 RPRQLLTSPKAVYVEAKYKQRIQIDYANLIGISVSSLSDSLFLVHVRQEDNKKQGDVYLQ 976
DB 901 RPRQLLTSPKAVYVEAKYKQRIQIDYANLIGISVSSLSDSLFLVHVRQEDNKKQGDVYLQ 960
OY 977 SDHYIETLTALTSADRVNININOGSITFAGGGRGDIIDTSGSELLTTKKNGLAV 1036
DB 961 SDHYIETLTALTSADRVNININOGSITFAGGGRGDIIDTSGSELLTTKKNGLAV 1020
OY 1037 VAPRLNSR 1044
DB 1021 VAPRLNSR 1028

```

```

RESULT 3
MYIB_DROME STANDARD; PRT; 1026 AA.
ID MYIB_DROME
AC Q23979; Q9W0H0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Myosin IB (MIB) (Brush border myosin IB) (BBMIB).

```

GN MYO61F OR CG9155.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mendibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
RC TISSUE-Head:
RX MEDLINE=94260541; PubMed=8201616;
RT "Morgan N.S., Skovronsky D.M., Artavanis-Tsakonas S., Mooseker M.S.;
RT "The molecular cloning and characterization of Drosophila melanogaster
RT myosin-1A and myosin-1B.";
RL J. Mol. Biol. 239:347-356(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RT "Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RT "Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RT "George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RT "Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RT "Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RT "Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RT "Bailey J.F., Agayanni A., An H.-J., Andrews-Efamikoch C., Baldwin D.,
RT "Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RT "Borkova D., Botchan M.R., Bouck J.P., Brokstein P., Brothier P.,
RT "Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RT "Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RT "de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RT "Dodon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RT "Dustin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RT "Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RT "Floodk A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RT "Harris N.L., Harvey D., Hewland T.J., Hernandez J.R., Houck J.,
RT "Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RT "Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RT "Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RT "Laoko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RT "Lin X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RT "Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RT "Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RT "Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclel J.M.,
RT "Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RT "Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RT "Shue B.C., Silden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RT "Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RT "Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RT "Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RT "Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RT "Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RT "Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu K., Smith H.O.,
RT "Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.";
RL "The genome sequence of Drosophila melanogaster.";
RN [3]
RP SEQUENCE OF 338-697 FROM N.A.
RX MEDLINE=94029356; PubMed=821225;
RT "Cheney C.M., Kravitt N.G., Verbsky J.W.";
RL "A new myosin I gene in Drosophila.";
RN [4]
RP Biochem. Biophys. Res. Commun. 195:1280-1288(1993).
RN [5]
RP SEQUENCE OF 630-1026 FROM N.A.
RA Caggese C.;
RN Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
STAGE.
RX MEDLINE=96069904; PubMed=7589814;
RT "Morgan N.S., Heintzelman M.B., Mooseker M.S.";
RT "Characterization of myosin-1A and myosin-1B, two unconventional
RT myosins associated with the Drosophila brush border cytoskeleton.";

	RL	-dev.	bio1.i7z-51-vA(1993).	
	CC	-FUNCTION:	INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG ACTIN FILAMENTETS (BY SIMILARITY).	
	CC	-SUBCELLULAR LOCATION:	CYTOPLASMIC. PROTEIN SHIFTS FROM THE BASOLATERAL TO APICAL DOMAIN IN ENTEROCYTES AND FOLLICLE CELLS.	
	CC	TISSUE SPECIFICITY:	IN THE EMBRIO, EXPRESSED IN GASTRIC CARCIE.	
	CC	MIDGUT CELLS OF THE PTEROPERICULOUS, AND IN THE MID AND HINDGUT. IN MICROVILLI. ALSO EXPRESSED AT HIGH LEVELS IN FOLLICLE CELLS DURING OOOENESIS.		
	CC	-DEVELOPMENTAL STAGE:	EXPRESSION STARTS AT 8-12 HOURS EMBRYONIC DEVELOPEMENT, CONTNUES TO INCREASE UNTIL THIRD LARVAL INSTAR,	
	CC	DISAPPEARS IN PUPAE AND IS PRESENT AT A LOW LEVEL IN ADULTS.		
	CC	EXPRESSIOIN IN EMBROGENESTIS IS CORRELATED WITH THE FORMATION OF A BRUSH BORDER WITHIN THE ALIMENTARY CANAL.		
	CC	SIMILIARTY:	CONTRAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.	
	CC	-SIMILIARTY:	CONTRAINS 2 IO DOMAINS.	
	CC	-CAUTION:	Ref.2 sequence differs from that shown due to erroneous gene model prediction.	

	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-slb.ch/announce/or send an email to licenselsb@slb.ch).		
	CC	EMBL;	U07596; AAA19591.1; .	
	DR	EMBL;	AEO03471;AAF47477.1; ALT_SEQ.	
	DR	EMBL;	L13070; ? NOT ANNOTATED_CDS.	
	DR	EMBL;	AJ000879;CAA04367.1; .	
	DR	HSSP;	P08789; IAMD.	
	DR	Flybase;	FBN0010246; Myo61F.	
	DR	InterPro;	IPIR000048; IQ_region.	
	DR	Interpro;	IPIR01609; myosin_head.	
	DR	pfam;	PF000063; myosin_head; 1.	
	DR	pfam;	PF00612; IQ; 3.	
	DR	PRINTS;	PR00193; MYOSINHNAVY.	
	DR	PRODcom;	PD000355; myosin_head; 1.	
	DR	SMART;	SMO0015; IQ; 1.	
	DR	SMART;	SMO0242; MYSC; 1.	
	DR	PSORTITE;	P550096; IQ; 1.	
	KW	Myosin;	Act-binding; Calmodulin-binding; Actin binding; Repeat.	
	FT	DOMAIN	? 1	
	FT	DOMAIN	? ?	
	FT	DOMAIN	697 724	
	FT	NIP BIND	105 112	
	FT	DOMAIN	583 590	
	FT	CONFICT	64	
	FT	CONFICT	221 221	
	FT	CONFICT	242 242	
	FT	CONFICT	262 262	
	FT	CONFICT	345 345	
	FT	CONFICT	551 552	
	FT	CONFICT	688 688	
	SEQE	SEQUENCE	1026 AA; 117955 MW; 60852B647FE98057 CRC64;	
Oy		Query Match	48.6%; Score 2626; DB 1; Length 1026;	
Dd		Best Local Similarity	51.5%; Pred No. 4.3e-159;	
Oy		Matches 532; Conservative 177; Mismatches 301; Indels 24; Gaps 8		
Oy		17 MESALTARDVGVDQPVLLLENFTSEAAFIETLRFRRENLITTYIGPVIVSYNPTDLOI	76	
Dd		1 METGLHERIRAGVDPPVLLENTYOSEBAFIGNKRRFOEDLITYYIQGVATSVNPYKOLPI	60	
Oy		77 YSRQHMERKRGSFYEVPHLTAVADTVYRALRTERDOAWMSGESGGATEAKRILO	136	
Dd		61 YYDDHVKAKVRNHKEFMPPHFIAVDNNARSLILENRGCVCVLISGESSGSKTAASKVIQ	120	
Oy		137 FYAETCPAPERGAAYDRLQLSGVPYLEAFGNKTLTNNDSSRFEGXYMDVFDFKGAPYG	196	

```

Db 121 FIAAGSGNQTVEGVKDKLLKSNPVLAFGNAKTNRNDNSRFGKMDIQDFKGAFIGG 180
Qy 197 HILSYLLEKSRVYVHONGHGERNFVVFQLEGGEEFLRLGLERNPOSTLYLVKGCANV 256
Db 181 NINLYLLEKSRVYVHONGHGERNFHIFQLLAGADEALLOELRERLADFTSYLTDGNGTV 240
Qy 257 SSINDKDMVKMRKALSVIDEFEDLSTIASVLHGNHFEADESNMQVTTENDL 316
Db 241 TRINDADSPKVOQALVIDFTKEOREIFGIYASLHNGNGF-TEVGNKAKVNSRDV 299
Qy 317 KYTRLGVEGTTLRALTRHRTIAKGEELLSPNLQAAVADALAKAVSRFTWLR 376
Db 300 VTAARLLGVASLELAALTRITDARGDVATSPNLQELAIYARADALAKAVYRLSFWLQ 359
Qy 377 KIRSLASDAESPSRSTTVGLDLYGFEPVQNHSPFQFCINYNENKLOLFELTLK 436
Db 360 RLWISLQAKETRA---SRNNVGLDLYGFELFQKNSPFQFCINENKLOLFELTLK 416
Qy 437 SEOEYEAAGIAMEPVQYFNKKIICDLVEEKEGIIISIDECLAPGEATDLFLEKLD 496
Db 417 SEOEYERBGIEIPIVEYFDNVIICLIEKKHGIISIDECLAPGEPTDKTFLEKLD 476
Qy 497 TVKPRPHFLTKLAODKTRKSLDRGFRLHLAGEVTVYVGLDKNDLFRNLKETWC 556
Db 477 KLAQHNYVYCHERAPAHIKKIMLDEFRLVHAGEVTVYVNGFLDKNDLFRDLKETLS 536
Qy 557 SSMNPIMAOCPDSEKSDKKRPETVATOKMSLDLVELLRKEPAVYICIRPNDAKOG 616
Db 537 KAGNGIVRSCFPEKELRSKREPETAITQIRASINNMIDLMCKESBYICIRPNDAKOG 596
Qy 617 RDEVLIRHGVYKYLGLMENLRYRAGFAVRRKYEAFQKYSLCETPMFM---AGRPDQ 674
Db 597 VENDELVLHQVYKYLGLMENLRYRAGFAVRRKYEAFQKYSLCETPMFM---AGRPDQ 656
Qy 675 VAVLVNHLGKVEBEYKMGRTKIFIRPKTLFTEDESLEYRROSATKIOANRGFHWK 734
Db 657 VOQLVNDLGMDEKRVGKTKELIRMPRLFTDEYAYOKKEHIAIIOAHKKGMRK 716
Qy 735 FLVKRKAICIOSMWMGTGRRAKRAKMAOTIRIRLNGFLRSPRCPENAFPLDHR 794
Db 717 YLTKRNOVILIMOSYCRKLAQAQAKRRRAKIRAFIKGFTTRNDAPNGENEEFTANK 776
Qy 795 ASFLNLRQDPRNVLDTISWTPRALPASELLELCNMVWKYCRISPEWQALQ 854
Db 777 RMLLRLAKELPTKYLDKSPHAPGCEASGILRHHLRLARILKLPQQRKREL 836
Qy 855 KAVASEIFKGGKNDYPOSYRLEFISRLTETI---SPRVLOSLSSEPQYAVPVYKTR 911
Db 837 KYLAERVFEGKKNNAASSVSTWFOEDRLPKHEHIOQVNDVASTFSEDLKQSFCTKDR 896
Qy 912 KGYKPRPOLLLTPSAVNVLEDKAVKQ-----RIDYANLIGISVSSLSDSLFLVHQR 964
Db 897 HGKYSRDRILLSSNAIYVLDGKTYKOKHRLPLDKIDF-----TLTNHNDLAVIRPL 950
Qy 965 ENNKOKGVYLOSADVHETLTALASDRVNNINI-NQGSITFAGPGRDGIIDTSSGE 1023
Db 951 DUKKRGDILILIPRIESTYIIDTVGTAIVSIVDRNLSLHNHVKGGVYIDIQGAE 1010
Qy 1024 LLITKAKNGHLAVY 1037
Db 1011 PGVYRDK-GHLVIT 1023

```

```

RESULT 4
MYHL_BOVIN STANDARD: PRT: 1043 AA.
AC P10568: Q9NS38:
DT 01-JUL-1989 (rel. 11, Created)
DT 01-JUL-1989 (rel. 11, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Brush border myosin I (BBM-I) (BBM1) (Myosin I heavy chain) (MHC)
EN (Brush border 110-kDa protein).
GN MYHL.

```

```

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88033016; PubMed=3667594;
RA Hoshimaru M., Nakanishi S.;
RT "Identification of a new type of mammalian myosin heavy chain by
RT molecular cloning. Overlap of its mRNA with preprothymosin B
RL J. Biol. Chem. 262:14625-14632(1987).
RN [2]
RP SEQUENCE OF 1-789 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92268028; PubMed=1587791;
RA Kawakami H., Moriyoshi K., Utsami T., Nakanishi S.;
RT "Structural organization and expression of the gene for bovine myosin
RT I heavy chain."
RL J. Biochem. 111:302-309(1992).
CC -1- FUNCTION: INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL: J02819; AAA30658.1;
CC DR PIR: A29483; A29483.
CC DR HSSP: P08799; 1NMD.
CC DR InterPro: IPR000048; IQ_region.
CC DR InterPro: IPR001609; myosin_head.
CC DR Pfam: PF00063; myosin_head; 1.
CC DR PRINTS: PR00193; MYOSINHEAVY.
CC DR ProDom: PD000355; myosin_head; 1.
CC DR SMART: SM00015; IQ; 3.
CC DR SMART: SM00242; MYSC; 1.
CC DR SMART: PS00096; IQ; 3.
CC KW myosin; Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
CC MultiGene family.
CC FT DOMAIN 1 681 MYOSIN HEAD-LIKE.
CC FT DOMAIN 1 697 719 IQ 1.
CC FT DOMAIN 2 720 742 IQ 2.
CC FT DOMAIN 3 743 772 IQ 3.
CC FT NP_BIND 101 108 ATP (BY SIMILARITY).
CC FT DOMAIN 571 593 ACTIN-BINDING (POTENTIAL).
CC FT CONFLICT 708 708 MISSING (IN REF. 2).
CC SQ SEQUENCE 1043 AA; 118866 MW; 581BEP1EA508B3C CRC64;

```

```

Query Match 39.08; Score 2106.5; DB 1; Length 1043;
Best Local Similarity 44.08; Pred. No. 4.5e-126;
Matches 455; Conservative 185; Mismatches 330; Indels 65; Gaps 21;

```

```

Qy 27 VGVDFVLLNFSEAFIENLRFRRENILTYIGVYLVSNVPRYRLOIYSQHMER 86
Db 8 VGVDFVLLNFSEAFIENLRFRRENILTYIGVYLVSNVPRYRLOIYSQHMER 86
Qy 87 GVSEYEPVPHLFAVDVYRALRTERDQAVMISGESGAGKATKRLQFVETCPAPE 146
Db 67 DYTFELPKPIYALANAYSLDRDQCIILITGESGAGKATKRLQFVETCPAPE 126
Qy 147 RGAAYVDRLLQSNPVLAEFAGNAKTLRNDNSRFGKMDYQDFKGAIPVGHILSYLLEK 206
Db 127 QVNSYVKEOLLQSNPVLAEFAGNAKTLRNDNSRFGKMDYQDFKGAIPVGHILSYLLEK 186

```


QY	27	VGOVOPVLELNTSEAFIENLRRRREULITTYICPVLSVNPYDLOIYSOHMERYR	86
Db	15	IGVGMVLEBP--NEETFI DNKKRBDHEIYITTYIGSVISVNPYSLPISEPEKDEY	73
QY	87	GVSEFVPHLEAVADTVYRALRTEREDDAVMISGSGAGKTEATRLLQFAENCPARE	146
Db	74	NNFELSHITRLSDEARSLRDOODKDCIITGESGAKTEASLVSYAANVGKGA	133
QY	147	RCGAVRDLLOSNPVLEAFGNAKTLNDNDSRFKGYMOQFDEKAPVGGHILSYLLEKS	206
Db	134	EYNOVKEOQLQSNPVLEAFGNKKTYANDNDSRFKGYMDIEPDFKGPDLGVSINYLEKS	193
QY	207	RVVHONHGRNHFVFOLEEGEELRRLGIERNQSLTYLVKGCCAVSSINXSDMK	266
Db	194	RVVKOPRGGRNHFVYQLLSGASELLYKLAKTERDSRYNLT--SUDSAKVNQDDAANR	252
QY	267	VNRKALSYIDTEDEVEDLLSIASVLYLGNTHFADEDSN---AQVTENOLKLTURL	322
Db	253	TVRNMOIYGFEDHEAEVLEVAANVLKGNIEFREPRESRVNGDDESKINDKNELKICEL	312
QY	363	LGVEGTTREALTHHKRIIAKGBELLSPLNBEAARAOLAANVSRITVLYRKINBL	382
Db	313	TSIDOVLERASFETVEKQKSVTTLNVAOAYTAROLAANLXLSRLESMVNRINESI	372
QY	363	ASKDAESPWSRSTVGLDIGFVFOHNSFEQFCINCNENKLOOLFETELTKSPOEY	442
Db	373	KAQTKVRRK----VMGVLDIYGFELPEDNSPEQFIINCNENKLOIFETELTKEQEY	427
QY	443	EAEGVAMRPOVYFNKKIICDLVEEKFGIISLDEBCLRGEATDITLEKLEDVYKPH	502
Db	428	IREDEWTHIDFENNAIICDLLENTNGTLAMDECBCLPGTVDTEFTEKKNQVCATIQ	487
QY	503	HFLTH--KLADOCTKRSLDRGEFRLHAGEVTVSYVGLDNNDLFPNLKETWCSSNN	560
Db	488	HPESMKSCSRNLNTITLPHSCFRLOHAGAKVLYVEGVEDKXNDLLIYDLSQAMMKASH	547
QY	561	PIMACQFQKSELS--DKRPEVTAQFKMSLQVLEILKSEKPAVITRCIIPNDAKOPGR	618
Db	548	SLINSLEFEGNDAKVALRPPYAGSQFASVATVTLRNLQTKNPNTIRCIKIPMDKAAHIF	607
QY	619	DEVILRQVKTGLMELNLRVRPAGPARYRKYAFLOKSLCPENPMWAGRPQGVANL	678
Db	608	NESLVCHQIRLTGLLENVRVRAGAFKQAIPECPLETKYKMLCKQWPHMKGARSQVEVL	667
QY	679	VRLHGLPEEYKMGKTIIFRPKTLFATEDSLLEVROSATKTIQAMNGFHMROKFLRV	738
Db	668	FNELEIPIPEEHSFGSKFIIRNPRTLFOLEDLRKQRBOLATLQIKYRGWKCGRHFILM	727
QY	739	KRSATCIOSMWRGTIGRRKAKRKNAAOTIRLLINGF-----ILRHSRCPENA----	787
Db	728	KRSOVYIAMTRRYAQOKRQOYQISSLAVLOSITYRGWKAARKILRELKHQKRCEAATYTA	787
QY	788	-----FPLDH	792
Db	788	AYWHGTOARRELKRLKEEARNDANVAIWMATWLGKVRREYKFRFRAANGKIITYETLQR	847
QY	793	VRSFLLNLRLQOLR--NVLDTSWPPPALRBASELRELCKMKNW--KYCRSISPBM	848
Db	848	IVQYLTLEMKKMPSLSPIDKNMSPRLPLFDSTH--KELKRIFHLMCKCKYRDOFTQO	905
QY	849	KQOLQOKRVASEIFGKCDKNTPOSPYRLFISRLGTEELSP----RYLOSLSGEPTQVAV	904
Db	906	KLIYEERKEASELTKRKALKYPSVSGQFQAGYLT--EINKNPKKKLDALIEEKIITAE	962
QY	905	PVYVYDRKGYRPRPQLLTPSAVYVE--DAKYKORIDYANLTGISVSSLSDSLEFVLV	962
Db	963	VVNRINRANGSTSRIFELTNNNNLLADOKSGQISEVPLVDVTKVSSQNDGFEAVHL	1022
QY	963	QR--EDNKQOGVULSOHVITELTK--TALSABRVN--NIN-----NGOSITFA	1007
Db	1023	KEGESEAAKSGDFLSSDHLIMATIKLYTTLISOTKQKINIEISDEFLVQFRODKYCVFI	1082

Query Match	Best Local Similarity	Score 2080;	DB 1;	Length 1136;
Matches 463; Conservative 183; Mismatches 346; Indels 148; Gaps 19;	40.6%; Pred. No. 2,4e-144;			
DB	1083 GGNKNGSVP-----TCKRKNRLLEVA 1105			
RESULT 6				
MYIA_RAT	STANDARD:	PRT: 1136 AA.		
AC 005096:				
DT 16-OCT-2001 (rel. 40, Created)				
DT 16-OCT-2001 (rel. 40, Last sequence update)				
DT 16-OCT-2001 (rel. 40, Last annotation update)				
DE Myosin IA (Myosin I alpha) (MMI-alpha) (MMIA) (Myosin heavy chain myr 1).				
GN MYOIA.				
OS Rattus norvegicus (Rat).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Euteria; Rodentia; Sciurognath; Muridae; Murinae; Rattus.				
OX NCBI-TaxId=10116;				
RP (1)				
RC SEQUENCE FROM N.A.				
RC STRAIN=Sprague-Dawley;				
RX MEDLINE=93194945; PubMed=8449985;				
RA Ruppert C., Kroschewski R., Beehler M.;				
RT Identification, characterization and cloning of myr 1, a mammalian myosin-I.				
RL J. Cell Biol. 120:1393-1403(1993).				
CC -I- FUNCTION: MOTOR PROTEIN THAT MAY PARTICIPATE IN PROCESS CRITICAL TO NEURONAL DEVELOPMENT AND FUNCTION SUCH AS CELL MIGRATION, NERVE OUTGROWTH AND VESICULAR TRANSPORT (BY SIMILARITY).				
CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A (SHOWN HERE), B AND C; ARE PRODUCED BY ALTERNATIVE SPLICING.				
CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.				
CC -I- SIMILARITY: CONTAINS 5 IQ DOMAINS.				
CC -----				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sdb.ch/announce/).				
CC or send an email to license@isb-sdb.ch .				
CC -----				
DR EMBL; X68199; CAA84287.1; .				
DR HSSP; P08799; IMND.				
DR InterPro: IPR0000048; IQ_region.				
DR InterPro: IPR001609; myosin_head.				
DR Pfam; PF00063; myosin_head; 1.				
DR Pfam; PF00612; IQ; 6.				
DR PRINTS; PR00193; MYOSINHEAVY.				
DR Prodom; PD000355; myosin_head; 1.				
DR SMART; SM00015; IQ; 6.				
DR SMART; SM00242; MYSC; 1.				
DR PROSITE; PS50096; IQ; 5.				
KW Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat; Multigene family; Alternative splicing.				
FT DOMAIN 1 703 MYOSIN HEAD-LIKE.				
FT DOMAIN 1 704 729 IO 1.				
FT DOMAIN 730 750 IO 2.				
FT DOMAIN 750 778 IO 3.				
FT DOMAIN 780 807 IO 4.				
FT DOMAIN 808 837 IO 5.				
FT DOMAIN 837 866 IO 6.				
FT NP_BIND 108 115 ATP (POTENTIAL).				
FT NP_BIND 108 115 ACTIN-BINDING (POTENTIAL).				
FT DOMAIN 592 599 MISSING (IN ISOFORM 1B).				
FT VARSPLIC 794 823 MISSING (IN ISOFORM 1C).				
FT VARSPLIC 794 852 MISSING (IN ISOFORM 1C).				
SO SEQUENCE 1136 AA; 131917 MW; 69D6C2A84E9070E2 CRC64;				

```

27 VGVDFVLENTSEAFIENIRRRPRENLTITTYIGPVLSVNPYRDIQIYSROMEXR 86
15 IGVDFVLEPL-NEETFDLNLKRBDNEIYTYIGSVVIVSNPRSLPISPEKVEDYR 73
87 GVSFEVPHLFAVADTVYRALRTERDOAVMISGSGAGKEATRKLOFAETCPAPE 146
74 NNFIEFLSHIRALDEAVRSRLRDOCKOCLITGESGAGKEATRKLOFAETCPAPE 133
147 RGVAVDRLLQSNPVLEAFNAKTLRNDSSRFKGYMDVDFKAPVGGHILSYLLEKS 206
134 EYNOVEKLOLQSTPVLEAFNAKTVRNDSSRFKGYMDIEFDFKAPVGGHILSYLLEKS 193
207 RVYHONHGERNFHVPOLEGEETIRLGLERHPQSLYLVKQCKAVSSINCKSMK 266
194 RVYKOPGRERNHFVQLLSGASEELHLKLEPFSRNYL-SLDSAKVNVGDPAANR 252
267 VNRKALSVDFEDEDLISIVASYLHGINIFADAEDSN---AOVTENQOLYRL 322
253 TVRNMQIYGFSDPEASEVLEVAVALKGNIEFPESMNGLDSKIKDKNELKEICEL 312
323 LGVEGTLREALTNRKTIKGBELLSPLEBOAARALAAVSRFTWLVYRKNLSL 382
313 TSIDOVYLERAFSFRVEAKOKVSTLVVAQAYARALANKLSRLEFVNLNINSEI 372
383 ASKDAEPMRSTTVGLDIYGFVQHNSEFCINCNKELQOLFIETLKESEOEY 442
373 KQOTKVRK-----VAGVIDIYGFELFENSTFQOFTINCNKLOIFELTLKEOEY 427
443 EMEGIMPEVQYFNKTIODLVEKFKGIISLDEBCLRPGEATDLFLEKLEDTYKPPH 502
428 IREDIEMTHIDYFNNAIICDLIENNTNGILAMDECLRPGEATDLFLEKLEDTYKPPH 487
503 HELTH--KIADOKTRSLDGEPRILHYAGEVTVSTGVLKNDNLINLEKMETOSSN 560
488 HESRMSKCSRLINDTLTPHSCRIQHYAKVLYOEGVDVNDNLIDLSQAMKACH 547
561 PIMACOFDSELS--DKRPETVATQFKKSLQOLVELILSKSPAYTRICIPNDAKOPGRF 618
548 ALIKSLFPBGNNAKVALKRPPTAGSOFKSVATIMKNLQTKNPNYRICKPMDKAAHIF 607
619 DEVLIRHOVYVIGLIMENLVRAGVYRKRKYEAFLQRYSLCPEYTPMAAGRODGVAVL 678
608 SESLYCHOIRYIGLLENVRRAGVAFROAYEPCLEKRYMLKOTPMHKKGAPKASVEVL 667
679 VRHLGKPEYKMGKRTIFIRPKTIFAEEDSLVROSLARQIAOMGFHROKFLAV 738
668 FNELETPVEYRSGSKIFIRNPRTIFQLEDLRKQLELATILQYIGMKCRHFLIM 727
739 KRSALCIQSWMGRTIGRRKRAKRWAAQTRIRLNGF-----ILRHSRCPENA--- 787
728 KRSQVVIAMVRYRQOQKRYQIKSSALVQYIRGMKARKILRELKHOKCKEATTTIA 787
788 ----- 787
788 AYWHGTQARKERRRLKDEARNKHALVIAFWLGSFARRELKLEAKRKNANAVIWAY 847
788 -----FLIDHVRASFILNLRROJPR-NVLDTSMPTEPPALR 822
848 WLGAKYRREYKFFRANACKKIYEFLOHIVQYKLEEMKNKMSLSPIDKNMSRYLFL 907
823 EASELLRELCKMNMV---KYCRSISPEMKOOLQKAVASEIFPKGKNYPOSVPLFIS 879
908 DSH--KELKRIPIHWRCKKYPDOQKLIYEERKLESEIFKDKKALVPSVSGOPFOG 965
880 TRIGTEIISP---RVLOSGLSPIDQYAVPVYKDGKGVPRRROLLPSPAVIYE--D 933
966 AYI---EINKNPKYKLLDAIEKIIIAEVYKINANGKSVSRIFLTLNNLNLADQKS 1022
934 AKYKORIDYANLTGISVSLDLSFLYVOR--EDNKRQGDVVLQSDHVIETLTK--TAL 989
1023 GQIKSEYPLVDYKVMSSONGFVHLKESSEASKODFLFSSHLLIEMATKLYRTL 1082
990 SADRVN-NINI-----NOGSITFAGGPGRDIIDFTSGSELLITKAKNGLAVYA 1038

```

```

DB 1083 SQTOKNIEISDFLVQFRDQKVCYKFIQGNKNSVP-----TCKRKNRLLEVA 1134
RESULT 7
MYHL_HUMAN STANDARD; PRT; 1043 AA.
ID MYHL_HUMAN
AC Q9UBC5; Q9UBD7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brush border myosin I (BBM-I) (Myosin I heavy chain) (MHC).
GN MYHL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RA Birkle D.D., Munson S.J.;
RT "Human brush border myosin I";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Jejunal brush border;
RX MEDLINE=20119970; PubMed=10653589;
RA Li W., Wang J., Coluccio L.M., Matsudaire P., Grand R.J.;
RT "Brush border myosin I (BBMI): a basally localized transcript in human
RL jejunal enterocytes.";
RN J. Histochem. Cytochem. 48:89-94(2000).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99073867; PubMed=9858156;
RA Skovron J.F., Bement W.M., Mosseker M.S.;
RT "Human brush border myosin-I and myosin-Ic expression in human
RL intestine and Caco-2Bb cells.";
CC Cell Motil. Cytoskeleton 41:308-324(1998).
-1- FUNCTION: INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG
CC ACTIN FILAMENTS (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AF127026; AAD31189.1; -
CC EMBL: AF105424; AAC78645.1; -
CC EMBL: AF009961; AAC27437.1; -
CC HSSP: P08799; 1AMD.
CC Genew; HGNC:7595; MYO1A.
CC InterPro; IPR001048; IQ_region.
CC InterPro; IPR001609; myosin_head.
CC Pfam; PF00612; myosin_head; 1.
CC Pfam; PF00612; IQ; 3.
CC PRINTS; PR00193; MYOSINHEAVY.
CC Prodom; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 3.
CC SMART; SM00242; MYSC; 1.
CC PROSITE; PS00096; IQ; 2.
KW Myosin; Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
KW Multigene family.
KW DOMAIN 1 681 MYOSIN HEAD-LIKE.
FT DOMAIN 1 681
FT DOMAIN 697 719 IQ 1.
FT DOMAIN 720 742 IQ 2.
FT DOMAIN 743 772 IQ 3.
FT NP_BIND 101 108 ATP (POTENTIAL).
FT ACTIN-BINDING (POTENTIAL).

```

FT CONFLICT 446 446 Q -> P (IN REF. 3).
 FT CONFLICT 712 712 R -> G (IN REF. 3).
 SQ SEQUENCE 1043 AA; 118400 MM; EESD79D9DE85DC CRC64;

Query Match 38.28; Score 2064.5; DB 1; Length 1043;
 Best Local Similarity 44.5%; Pred. No. 2.1e-123;
 Matches 440; Conservative 186; Mismatches 320; Indels 43; Gaps 19;

27 VGVQDFVLENFSTSEAFETENLRREMLITYITGPVLSVNPYDLOIYSRQHERR 86
 8 VGVEDVLEPLVEE-SLKNLOLRKREKKEITYITGVISVNPYQDLIYSPREFLAKY 66
 87 GVSFEVPHLFAVDYVYRALTERDQAVMISGSGAGKTEATKRLDQFAETCPAPE 146
 67 DYFVELKPHIYALAVAVOSLRDRDQCILITGSSGSKTEASLVMSYAAVCGKE 126
 147 RCGAVDRLLQSPVLEAFNGAKTLNDSNRFKYMVDQFPGAPVGGHLLSYLEKS 206
 127 QVNSVEKOLQSPVLEAFNGAKTIRNNNSRFGKYMDEDFKSGPLGVITNLYLEKS 186
 207 RVVHONGERNPHVFQOLLEGGEETLRRLGLENRQSLYLKGCQAVSSINDKSPK 266
 187 RLKQKLGKERNHIFQOLAGADEQLKALKLERDTGTAYL-NHEVSHVDMDDASSER 245
 267 VMKALSVIDEDEVEDLLSIYASVYLHGNIFPADEDSNAQVTE-----NOLKYLTR 321
 246 AVQSAVAIVGSEEEIROYLEVTSWLKLGAV-LVADEFOAGSIPASGIRDRGVAEIGE 304
 322 LGCEVETLREALTHKTIKAGEBELLSPNLDOAAVADALAKAVYSRTFTWLVRKINS 381
 305 MGLNSEEVEERLACSTMETAKEKVYALNVQAQYARDAKLNKITSRLFDWIVRINES 364
 382 LASKDAPSMRSTVYGLDLDITGFVFNQNSPEOCINVCNEKLOQLEITELTSEDEE 441
 365 IKVGIEKK-----VMGVLDITGFIELEDNSFEOGVINVCNEKLOQLEITELTSEDEE 419
 442 YEAEGLAMEPVQYFNKKIICDLVEEKFGIISILDECLRPBETDLPFLKLEDTYKPH 501
 420 YKREGIPMTKVDYFDNGIICKLLEHNOGILAMLEDECLRPVSDSPFLANKLQFSKH 479
 502 PHFLTHKLADQTKSLDR-----GERLLHAGEVYTYSTGLDKNNDLFLFNLETKMOS 557
 480 GHT--ESKVTQNAQROYDHTMGLSCFRICHYAGKYTYNTSIDKNNDLFLFDLLQAMWK 537
 558 SMNPWAOCEFDKS--ELSDKKRPETVATQFKMSLLQVLVEILSKPEVATRCIKPDAPK 615
 538 AQHPILRLSFPPEGNPKQASLKRPPAGAFKSSVALIMKNLISKBNYIKCIKPEHQR 597
 616 GRDEYLLRHQVYVGLMELNVRRAGFAYRRKYEAFLQRYKSLCPETPMWAGRPDQV 675
 598 GOFSSDVLVATQARYIGLENNVRRAGYAHROGYGPFLEERYLLSRTPMNGDREGV 657
 676 AVLVRHLGYKPREYKMKRKTIFIRPKTIFATEDSLEVARQSLAKRQIQAAMGFIMROKE 735
 658 EKVVLGELSSGELAFGKTKIFIRSPKTLFYLEEORRLQLQDLALIKIYGMGRCHY 717
 736 LNVKRSALCIQSMNRGTGLRRKAARKMAAOTIRRLIRGLRHSPP--CPENAFPL-D 791
 718 QLMRSQILISSMFGNMOKKCYGKITKASVLLQAFVGRMKARKKRYKTFSEALTLIAD 777
 792 HVRAS---FLNLNRQLP-RNVLDTSWETPP-PALREASSELLRELCKMKNWYKCR-- 842
 778 FTYKSWVQKFLGLTKNNLPSTNVLDKTPAAYKCLSTPANOQLQLEFYQ--WKCKRPD 834
 843 SISPEMKOOLQKAAVASEIFEGKKKNYPOSVRRLISTRLGTEESLPRVLOSLSGE--PI 900
 835 QLSPKOVELRKELKCLASELFGKKKASYPQSVIPFCGDYIGLG--NPKLQKLKLGEEGPV 893
 901 QYAVPVVYDKRGYPRPRQLDITSASAVYEDAKVQR--IDVNLIGISVSDSLF 958
 894 LMAEAVKAVKVRNGNGTSSRILLTLTGHYVILDTTKKSQAKIYIGLDNVAGSVTSLKDLF 953
 955 VLVHOREDN-KQKGVVLOS DHVITLTK 986

Db 954 SIHLEMSVSGSGDFLVSEHVIETLTK 982

RESULT 8
 MYHL-CHICK
 AC P47807: 090573; STANDARD; PRT: 1045 AA.
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Brush border myosin I (BBM-I) (BBMI) (Myosin I heavy chain) (MHC).
 GN MYHL.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96114119; PubMed=8681398;
 RA Collins K., Matsudaira P.T.;
 RT "Recombinant expression of the brush border myosin I heavy chain.";
 RL Cell Motil. Cytoskeleton 32:151-161(1995).
 [2]
 RP SEQUENCE OF 46-1045 FROM N.A. AND SEQUENCE OF 332-343 AND 490-503.
 RX TISSUE=Intestine;
 RC MEDLINE=90078325; PubMed=2687288;
 RA Garcia A., Coudrier E., Carboni J., Anderson J., Vandekerckhove J.,
 RA Moosker M., Louvard D., Arpin M.;
 RT "Partial deduced sequence of the 110-kD-calmodulin complex of the
 RT avian intestinal microvillus shows that this mechanoenzyme is a
 RT member of the myosin I family.";
 RL J. Cell Biol. 109:2895-2903(1989).
 [3]
 RP 3D-STRUCTURE MODELING OF 699-731.
 RX MEDLINE=97148341; PubMed=8894973;
 RA House A., Silver M., Cohen C.;
 RT "A model of Ca(2+)-free calmodulin binding to unconventional myosins
 RT reveals how calmodulin acts as a regulatory switch.";
 RL Structure 4:1475-1490(1996).
 CC -1- FUNCTION: COULD PLAY AN IMPORTANT ROLE IN MORPHOGENESIS AND
 CC FUNCTION OF INTESTINAL MICROVILLI.
 CC -1- TISSUE SPECIFICITY: INTESTINE.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 CC EMBL: U04049; AAB38373.1;
 CC EMBL: X58479; CA41388.1;
 CC PIR: A33620; A33620.
 CC PDB: 1AUI; 17-SEP-97.
 CC InterPro: IPR000048; IQ_region.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF00063; myosin_head.1.
 CC Pfam: PF00612; IQ.3.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC ProDom: PD000353; myosin_head.1.
 CC SMART: SM00015; IQ.5.
 CC SMART: SM00242; MYSC.1.
 CC PROSITE: PS50096; IQ.3.
 CC MYOSIN: Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
 CC Multigene family; 3d-structure.
 CC DOMAIN 1 684 MYOSIN HEAD-LIKE.
 FT DOMAIN 701 727 IQ 1.
 FT DOMAIN 723 750 IQ 2.

Matches 406; Conservative 177; Mismatches 292; Indels 32; Gaps 12;

QY 27 VGVDFVLENTSEAFIENLRRENTLITYIGVAVSVNPRDQIYSQHMERYR 86
 DB 8 VGVDFVLENTSEAFIENLRRENTLITYIGVAVSVNPRDQIYSQHMERYR 86
 QY 87 GVSFEVPHLFAVADTVYRALRTERDQAVMISGSGAGKTATKLLQFVETCPAPE 146
 DB 67 DYFEYELKPHIYALANAVASIKDRDQOCILITGESGAKTKXXXKLMSYVAAGVCGE 126
 QY 147 RGAVNRDLQSNPVLEAFGNNAKTLRNDSSRFKGYMDQDFKAPVGGHLSYLEKS 206
 DB 127 QVNSVKEBOLQSNPVLEAFGNNAKTLRNDSSRFKGYMDQDFKAPVGGHLSYLEKS 186
 QY 207 RVVHONGERNPHVYVQLLEGSEETLRRLGLERNQSYLYLKGCACAKYSINDKSDMK 266
 DB 187 RVVHONGERNPHVYVQLLEGSEETLRRLGLERNQSYLYLKGCACAKYSINDKSDMK 245
 QY 267 VMKALSVIDFTEDEVEDLSIVASVHLGNHFAADESDN---AQVTENQKYLTRL 322
 DB 246 AVQHANSVIGFSEETIRNOVLEVALYKLGNAVLTDFEQANGIPASGIDGKIOIETGEM 305
 QY 323 LGVEGTLRALTHRIKANGELLSPLENOAAYADALAKAVYRTFTWVYKRNRSI 382
 DB 306 MGLNTELEALCSRTMETGKEVYVYLVANTQAVARDALAKAVYRTFTWVYKRNRSI 365
 QY 383 ASKDAESPWRSTTVGLDIDYGFYFQHNSEFQFCINCNKLOOLFTLTKSROEY 442
 DB 366 KVGTGEKK---VGVLDIYGFELLEDNSFQFVLANCNELQOVFTELITKEOEY 420
 QY 443 EAGCIAMEPVQYNNKICDIVERKGIISIDECLRGAEATDLTFLEKEDTVKPPH 502
 DB 421 KRGIPMTKVEYEDNIGLHSHQSGILAMDECLRGVYSDSTFLAKLQNFSKHS 480
 QY 503 HFLTHKLAADOKTKRSDDR---GEFRLHAGVTVSYVGLDKNDLLEFRNLKEMCS 558
 DB 481 HY---ESKVSOMARQYDRTMGISCFTISHAGVTVYVNGFIDKNDLLEFRNLKEMCS 538
 QY 559 MNPIMACDSELSDK--KRPETVATQFKMSLQLELIRSEKPAVYRCIRPNDAKOPG 616
 DB 539 OHDLKSLFPEGNPKFASLKRPTACTQFNKSNVAVMLKMLNYSNPNYICIRPNDOQNG 598
 QY 617 RPEDEVILRHQVTKLGLMENIRVARRACAFYRKYEAFLQYKSLCPEPTWMAAGROPDVA 676
 DB 599 RFTSEVMVQARLYGLENVYRVARAGYAFROGKPELERYRLRSTWPMRMDDEGVE 658
 QY 677 VLVRLHGYKPEEYKMGRTKIFIRFPTLFATEDLSLEVRROSLATKIQAMRPFHMKRL 736
 DB 659 KVLGSITLSSEELAYKTKIFIRFPTLFATEDLSLEVRROSLATKIQAMRPFHMKRL 718
 QY 737 RVKRSALICIQSMWRGTLGRKRAKMAQITRLIRGLFHLNHSRCPENA-----F 788
 DB 719 OMKRSQILISAMFRGKQKHYKIRSSVLLQAFRGARAKNRYKRSAGALTLANF 778
 QY 789 FLDDHVASFLNLRRLP--RNVLDTSWTPRP--PALREASELLRELCMKMMWKYCRSISP 846
 DB 779 IYOSMAQKFLNKKLPSTKVLNDNTWPAAPYRCFNTANQLORLFYOMKCKKFRLOSLP 838
 QY 847 EMKQOQQKAVASELFGKKNVPOSVPRILFISTRLTEISPRVLOSIGSE--PIQVAV 904
 DB 839 KOVQTLREKLASSELFGKKNVPOSVPRILFISTRLTEISPRVLOSIGSE--PIQVAV 897
 QY 905 PAVKXDR 911
 DB 898 TVKXVNR 904

RESULT 10
 MYHL_RAT
 ID MYHL_RAT STANDARD: PRT: 842 AA.
 AC 062774:
 DT 16-OCT-2001 (rel. 40, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)

DT 16-OCT-2001 (rel. 40, last annotation update)
 DE Brush border myosin I (BBM-1) (BBM1) (Myosin I heavy chain) (MHC)
 GN MYHL OR BBM1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN NCBI:TaxID=10116;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver, and Testis;
 RA MEDLINE=95298044; PubMed=7779104;
 RX Balish M.F., Coluccio L.M.,
 RT Identification of brush border myosin-I in liver and testis.";
 RL Biochem. Biophys. Res. Commun. 211:331-339(1995).
 CC -1- FUNCTION: INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG
 CC ACTIN FILAMENTS (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: U25148; AAA9132.1; -.
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR000048; IQ_region.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00612; IQ_3.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 3.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS0096; IQ; 2.
 DR PROSITE: PS0096; IQ; 2.
 DR MYOSIN; Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
 KW MultiGene family.
 FT NON-TER 1
 FT DOMAIN 1 673 MYOSIN HEAD-LIKE.
 FT DOMAIN 689 712 IQ 1.
 FT DOMAIN 713 733 IQ 2.
 FT DOMAIN 735 763 IQ 3.
 FT NON-TER 842
 SQ SEQUENCE 842 AA; 97210 MW; DDAFDCB66308316B CRC64;

Query Match 35.8%; Score 1932.5; DB 1; Length 842;
 Best Local Similarity 45.6%; Pred. No. 3.7e-115;
 Matches 388; Conservative 165; Mismatches 269; Indels 29; Gaps 10;

QY 28 VGVDFVLENTSEAFIENLRRENTLITYIGVAVSVNPRDQIYSQHMERYR 87
 DB 1 GVEDLILEPL--DEESLKNLQURKREKKEIYIGVAVSNPNPQOLPYGPEEIRATYRD 59
 QY 88 VSEFEVPHLFAVADTVYRALRTERDQAVMISGSGAGKTATKLLQFVETCPAPE 147
 DB 60 YTFELKPHIYALANAVASIKDRDQOCILITGESGAKTKXXXKLMSYVAAGVCGEQ 119
 QY 148 GGAVNRDLQSNPVLEAFGNNAKTLRNDSSRFKGYMDQDFKAPVGGHLSYLEKS 207
 DB 120 VNSVKEBOLQSNPVLEAFGNNAKTLRNDSSRFKGYMDQDFKAPVGGHLSYLEKS 179
 QY 208 RVVHONGERNPHVYVQLLEGSEETLRRLGLERNQSYLYLKGCACAKYSINDKSDMK 267
 DB 180 RVVHONGERNPHVYVQLLEGSEETLRRLGLERNQSYLYLKGCACAKYSINDKSDMK 248
 QY 268 VMKALSVIDFTEDEVEDLSIVASVHLGNHFAADESDN---AQVTENQKYLTRL 323
 DB 239 VQNMAYVIGFSEETIRNOVLEVALYKLGNAVLTDFEQANGIPASGIDGKIOIETGEM 298

OY 324 GVEGTTTREALTRKRIIANGCELLSPNLLEOAAVARDALAKAYSRFTWLVKIRNSIA 383
 DB 299 GLNSVLELRLKLSRTETGKEKAVYLVNIOAYADALAKNIYSLEFMTYKRIWIESIK 358
 OY 384 SKDAESPWSRSTVILGILDIYGEVFOHNSFEQFCINCEKIQLOQFIETLTKSEOEYE 443
 DB 359 VGTGE-----KRVKMGVLDIYGFELLEDNSFEQFVINCENKIQOYFIELTLKEOEYEK 413
 OY 444 AEGIAEPVOYFNNKIIICDLVERKEFGIISILDECLRGGEADLTFLEKLEETVPHH 503
 DB 414 REGIPMTKVEYFNGIICDLIEHSORGILAMDEECLRGVSDTFILAKNLQJESKSH 473
 OY 504 FLTHKLADOKTRSLR---GEFRILHAGEVYVGTGLDKNDLFRNLKETKSCSSM 559
 DB 474 Y--ESKYTONAOROYRDTMGLSCFRICHAGKTYVTSIDKNDLRLDLSQTMWKKO 531
 OY 560 NPIMAOCEDKS--ELSDKRPPEVAFQFKMSLQLOVEILRSKREPATRICIKPDAPQGR 617
 DB 532 HPLHLSLFPRGNPKESAPKRPAGTOFKNSVALINKNLYSKNPNYIRICIKPDQOOG 591
 OY 618 FDEVLIHQYKTYGIMENILVRRAGFAYRRKYEAFIORYKSLCPETMPMAGRPDQVAV 677
 DB 592 FTSELVWQARYIGLEENVRAGAFRAOAYFLEKRLILSRSTWPRNGEDREGVEK 651
 OY 678 LVHILAGKPEEYKMGRTKIFIRPKTLFATEDSLLEVRSOLATKIOAMRGFHWKQFLR 737
 DB 652 VLCSLIISSEELAYGRKIFIRPKTLFYLEEGRRLRLQOLATLQKVGWGRCPHYOO 711
 OY 738 VKNSAICIGSMWRTGLGRKRAAKRMAQOTIRLNGF-----ILRSPCEPENAIF 789
 DB 712 MKRSQILISAMFRGNKOKKHYKIRSSVLLIOAFVRSWKRKKRYRFRSGARITLANF 771
 OY 790 LDHVRASELNLRRLOP--RNVLDTSWPTP--PALREASELREICKMMNWKCRSISPE 847
 DB 772 YOSTOKFLNLKKNLPSTVLDNTWPAFRCPTNANQELHFLQWCKKTRIDOLSPK 831
 OY 848 WKQOLOQKAVA 858
 DB 832 QVOTLRKICA 842
 RESULT 11
 MYIA.DROME STANDARD; PRT; 1011 AA.
 AC 023978;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE MYOSIN IA (MIA) (Brush border myosin IA) (BBMIA).
 GN MYO31DF OR CG7438.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
 RC TISSUE=Embryo;
 RX MEDLINE=94260541; PubMed=8201616;
 RA Morgan N.S., Skovronsky D.M., Artavanis-Tsakonas S., Mooseker M.S.;
 RT "The molecular cloning and characterization of Drosophila melanogaster
 myosin-IA and myosin-IB."
 RL J. Mol. Biol. 239:347-356(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang O., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Abgayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brockstein P., Brötter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Daventrop L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasio P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spiteras R., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 [3]
 RP FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
 STAGE.
 RX MEDLINE=96069904; PubMed=7589814;
 RA Morgan N.S., Heinzelman M.B., Mooseker M.S.;
 RT "Characterization of myosin-IA and myosin-IB, two unconventional
 myosins associated with the Drosophila brush border cytoskeleton."
 RL Dev. Biol. 172:51-71(1995).
 CC -1- FUNCTION: INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG
 ACTIN FILAMENTS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. PROTEIN SHIFTS FROM THE
 BASOLATERAL TO APICAL DOMAIN IN ENTEROCYTES AND FOLLICLE
 CELLS.
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED IN GASTRIC CAECAE,
 MIDGUT CELLS OF THE PROVENTRICULUS, AND IN THE MID AND HINDGUT. IN
 THE LARVAL GUT BRUSH BORDER, EXPRESSION IS IN THE TERMINAL WEB
 DOMAIN. IN THE ADULT GUT BRUSH BORDER, EXPRESSION REMAINS IN THE
 WEB DOMAIN AND HAS ALSO MOVED INTO THE MICROVILLI. ALSO EXPRESSED
 AT LOW LEVELS IN FOLLICLE CELLS DURING OOGENESIS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY
 THROUGHOUT DEVELOPMENT TO ADULTHOOD WITH HIGHEST LEVELS AT THE END
 OF LARVAL DEVELOPMENT. EXPRESSION IN EMBRYOGENESIS IS CORRELATED
 WITH THE FORMATION OF A BRUSH BORDER WITHIN THE ALIMENTARY
 CANAL.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 TO DOMAINS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL: U07595; AAA19590.1; -
 DR EMBL: AE003628; AAF52966.1; -
 DR HSSP: P08799; IAMD.
 DR FlyBase: FBgn0011673; MYO31DF.
 DR InterPro: IPR000048; IO_region.
 DR InterPro: IPR001609; myosin_head.

DR pfam: PF00063; myosin_head; 1.
 DR pfam: PF00612; IQ; 2.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD00035; myosin_head; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS00096; IQ; 1.
 DR Myosin: ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
 KW Multigene family.
 FT DOMAIN 1 694 MYOSIN HEAD-LIKE.
 FT DOMAIN 695 721 IQ 1.
 FT DOMAIN 721 736 IQ 2.
 FT NP_BIND 100 107 ATP (BY SIMILARITY).
 FT DOMAIN 567 589 ACTIN-BINDING (BY SIMILARITY).
 FT SEQUENCE 1011 AA; 117094 MW; 3120606585D50FA CRC64;
 Query Match 33.4%; Score 1800.5; DB 1; Length 1011;
 Best Local Similarity 39.6%; Pred. No. 1.2e-106;
 Matches 426; Conservative 169; Mismatches 341; Indels 139; Gaps 29;

26 GVQDFVLENTSEAFLENTLRRLRRRLTYTGYVAVNPYRDLOIYRQHMERYG 87
 8 GVQDFVLENTSEAFLENTLRRLRRRLTYTGYVAVNPYRDLOIYRQHMERYG 66
 88 VSFYEPPLFAVDVYRALRTERDQAVMISGSGAGTETATKLFQYAECPAP-- 145
 67 RELFENAPHLFAIDASAVRLKQRODCLISGSGAGTETATKLFQYAECPAP-- 126
 146 ---ERGGAVDRRLQSNPVLFAFGNAKTLRNDSSREGKWDVDFEFGAVGHHISYL 202
 127 NEIER---VKVYLIOSNAILETFGNAKTRNDSSREGKWDVDFEFGAVGHHISYL 183
 203 LEKSNVHONGERNERFHYFOLLGEGEETLRRLGLENPQSYLYLVGQCAKYSINDK 262
 184 LEKSNVHONGERNERFHYFOLLGEGEETLRRLGLENPQSYLYLVGQCAKYSINDK 240
 263 SDMKVRRKALVIDTEDEVEDLSIVASVHLGNIHNAEDSDNAQYTTENQKYLRL 322
 241 SDYKGTCAKFTLGFSTDEQVITMTAAVHLGVEQTLDE-ELVSNQNHKSTKRL 299
 323 LGVETTLREALTHKTIKAEELISPLNQAAYARALAKAVYSKRTFWLVKINSRL 382
 300 LQVETTELSTALTRKVIAGNVQKDMATQAEKGAALAKAYDRFTLTIIRINAI 359
 383 ASKDAESPMSSTYVGLDLYGFEVFOHNSFEQFCINCEKIOQLTELTISEQEE 442
 360 LFRGSKTQA-RKNSVIGVLDLYGFEVFOHNSFEQFCINCEKIOQLTELTISEQEE 418
 443 EAEGIAWEPVOYFNKKIICDLYVEKFKGISILDECLRPGEATDLEKLEDTYKRP 502
 419 QREGIEMTNIEFYFNKKIICDLYVEKFKGISILDECLRPGEATDLEKLEDTYKRP 478
 503 HPLTHKLAOKTRSL-DRGEFRLHVAEYTVYGTGLDNNLLFR-----NLKET 554
 479 HTSQQ--KPLDKELKREDFRTHAGDVYININGTEKNTKTLQDFERRLLHNSDA 536
 555 MCSNNPIMAOFCPSKSELDKRPETVATQFMSLQVLETLRSEKPAVYICIKPNDAK 614
 537 NLSEKMPRGADIKT-----TKRPLTACTLFQRMADLVYLLKKEPPYVACIKPNDAK 592
 615 PGRDEVILRIQVYKGLGMLNLRVRRAGFAVRKRYEAFLOKYKSLCPETWPM-AGRPD 673
 593 STVEDEEVEHQVRLGLENLVRVRRAGFAVRKRYEAFLOKYKSLCPETWPM-AGRPD 652
 674 GVAIVYRLHLYKPEYKMGRTKIFRPKTLFATEDSLFVROSLATQAAAGFIRH 733
 653 GVRVLEEKTF-ADQVKGHTKIFTRSRITFALEHONENIPIHIVYLLQRRVGVYRR 711
 734 KFLVRSALICIQSWNRGTLGRRAAKRMAQTIRLLIRGFIIRSPRENAFLDHY 793
 712 NFKKK-----AAILTVRAVYKAVKLR----- 732
 794 RASFLNLRLQ--PRNVLD-----TSWTPPALREASSELLR-----ELCKMMV-WKYCRS 843

DB 733 --STVQELANLRKAKMRDYGSTQMPQPLGKRYEAKLHRMDFWRAMMLTHKPRS 790
 QY 844 ISPEMKOOLQOKAVASEIFKRDYPOSVAPRLPISRLTEISPLVLSGS----- 897
 DB 791 ---EW-PQLHQLQIATAALAGRRRYWGA--RRWGVYLANSDENSGEYVNSIKNIN 844
 QY 898 -----EPQYAVPVYKDRGKYPKRPOLLTPSAVAVVEDAK-----VKORIDYAN 944
 DB 845 HPADGETFOOVLESFVKKNHFN-KQANRAFIYSDSTIHKLGIKKFKDMKRTIKIRE 903
 QY 945 LTGTSVSLSDSLFVLHVQREDNKQGDVY--LQSDHVIETLTKTALSADRNV----- 995
 DB 904 LITISVSPGRQDLVFN-----SSKNK-DLVFSLESE-----TLPEKEDRIEYGVYIC 951
 QY 996 -----NININGSTIFAGPGRDEIIFTSGSELLITKANKGLAVAP 1039
 DB 952 KKYHDLGTLELVAVNTNISCRLDKARITVEASVVEVPNRPKGNITFEVP 1006

RESULT 12
 ID MYSE_DICDI STANDARD; PRT; 1003 AA.
 AC 003479;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Myosin IE heavy chain.
 GN MYOE OR DMIE.
 OS Dictyostellum discoideum (Slime mold).
 CC Eukaryota; Mycetozoa; Dictyostellidae; Dictyostellum.
 OX NCBI_Taxid=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX3;
 RX MEDLINE=93277957; Pubmed=8504170;
 RA Urrutia R., Jung G., Hammer J.A. III;
 RT "The Dictyostellum myosin IE heavy chain gene encodes a truncated isoform that lacks sequences corresponding to the actin binding site in the tail."
 RL Blochim. Biophys. Acta 1173:225-229(1993).
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE ACTIVITY THAT IS ACTIVATED BY ACTIN. MAY PLAY A ROLE IN MOVING MEMBRANES RELATIVE TO ACTIN.
 CC -1- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: L06805; AAA33201.1; -
 CC PIR: S33760; S33760.
 DR HSP: P08799; IAMD.
 DR DictyDb: DD01049; MYOE.
 DR InterPro: IPR000048; IQ_region.
 DR InterPro: IPR001609; myosin_head.
 DR pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD00035; myosin_head; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS00096; IQ; 2.
 KW Myosin; Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
 KW Multigene family.
 FT DOMAIN 1 693 MYOSIN HEAD-LIKE.

Query Match 32.28; Score 1739; DB 1; Length 1003;
 Best Local Similarity 38.9%; Pred. No. 9.5e-103;
 Matches 416; Conservative 171; Mismatches 328; Indels 154; Gaps 29;

FT DOMAIN 694 722 IO 1.
 FT DOMAIN 716 745 IO 2.
 FT NP BIND 796 1003 NON ALPHA-HELICAL, C-TERMINAL DOMAIN.
 FT DOMAIN 101 108 ATP (BY SIMILARITY).
 FT DOMAIN 556 630 ACTIN-BINDING.
 FT DOMAIN 858 958 MEMBRANE-BINDING (POTENTIAL).
 SQ SEQUENCE 1003 AA; 114852 MW; BE758BEC035766f CRC64;

28 GVDFVLENTFSEAPFENLRARRRENILTYTIGVIVSVNRYRLOIYSRQRMERYG 87
 9 GVDFVLENTFSEAPFENLRARRRENILTYTIGVIVSVNRYRLOIYSRQRMERYG 67
 88 VSEVEVPHLEAVADTVYRALTERRDQAVMISGESGAGTEATKRLQFY----AETCP 143
 68 RYKEMPPIYALANAYRSMRQSENOCVIISGESGAGTEASKRMQFLFVSSNOSP 127
 144 APERGAVRRLQSNPYLEAFGNATKLRDNRSGKTYMDVOPDKGAPVGGHLLSYLL 203
 128 NGRISKMLIR-EQSNPLLEAFGNATKLRDNRSGKTYMDVOPDKGAPVGGHLLSYLL 186
 204 EKSRYVHONGENRHFVYOLLEGGEEETLRGLERNPQSYLYLKOGCAKYSINDS 263
 187 EKSRYVGRTOGERSFHFYOMLGLSOSKINELGLIPNAPAYEYLKSGCFDYSTDIDSG 246
 264 DKRWKRALSVIDFTEDEVEDLISVAVLHGNHFADEDSNAQVT---ENOLKY 318
 247 EFKIYVAMETTLCKESDONSIRWILAIHININIFAEAEBOORTTGVKVSDTSLAA 306
 319 LTRULGVEGTLEALHRIK---TAKGELL-S-PLNBOAARARALKAAYSRFTETL 374
 307 AASCLTKDQSLTALCYRISGVGRKSVISVPMOCNAAASRLAALERTLPLNML 366
 375 VKRINRSKASDAEBSRSTVYGLDIYFEVFOHNSPEOPCINYECKELQOFLTEL 434
 367 VSKIN-----TLSTYHREGVYIGLDIYFEVFOHNSPEOPCINYECKELQOFLTEL 419
 435 LKSEOEYEAJAMEVOYFNKKIICDLVEEFKGIISLDECLRPGEATPLTLEKL 494
 420 LKSEOEY--KGIEMKIEYF-NKPCIELEKRPGLISLDECL-IAKSTQDTPLDST 475
 495 EDYVKKPHFPLTKLADQKTRKSLDGE--FRLLHYAGETVYVTEFLKNNDLPLNKL 552
 476 CKQEPKPHLOSYVS-----KDRSNDICFRLAKHTAGDYTVYVGRGLDKNKTPLGDL 530
 553 ETWCSSNPIMAOCFDSKSLSD-KKRPETVATQFKMSLQVLEILSKREPAYIRCIKPN 611
 531 SSMQSSDPLVQGLFPTPRPDSKRRPRTAGSOFRRAMNALLITLLACSPHYVRCIKSND 590
 612 AKQGRFDEVLIRHOKYVGLMNLVRRAGFAVRRKYEFLORYSLCPEYPMAGRP 671
 591 NKAGVIEDEDRHRQVYIGLLENVRRAGFAVRRKYEFLORYSLCPEYPMAGRP 650
 672 QDGAVALVRLHGLKPYEYKGRKIFIRFPTLFATEDSLSEVRSQSLATQIQAAMGFHW 731
 651 KQATELLQOHNDKEIRMGKTKVIRIPRTLFYEERKELMLPRTIVILQKTMGYSYR 710
 732 RQFLVYKRSALCISQMWGRTGLRRKAAKRWAAQOTIRRLNGF-ILRHSRCPENAFEL 790
 711 RSKW-NORKAAIKIQLFY-----RSRYRKNW---FREIHRAFKVDIARPOWQKQVF-- 757
 791 DHRASFLMLRROLPRNVLDTSMPRPRLRASELRLCKMKNVWYCRISF--EM 848
 758 -----WPKDPSILDRAYDLTKRI--HNCMRAEKMTISLGAG 791
 849 KQLOQKAVASEIFKGGK-----DNYPSVRLFTIRLSTETI 887
 792 QNHKROKVMYVDLFHGKKKMDFRHRHDADYLEKPMWNOCKYVLAQONLF----- 841
 888 SPRVLOSGLSEPIQYAVPVVKYDRKGYKPRROLITPSAVVIVEDAKYKOR--IDYAN 944

DB 842 -----STYGDTEVFLADYIVKPNKGV-PQRRGIYVTEGNTNKKDKPNKYNKMGTPLYD 895
 QY 945 LTGISVSSLSDFLVH---OREDNKOG-----DVLQSDHYETITLTKIALS 990
 DB 896 VRSISIPMADFLVHLHKAQPRDFLDLGGNGEYAVSEITVYVQO---VLKTLGVKLS 952
 QY 991 ADVNNININ-----OGSIT-----FAGGPRGRIIDE 1018
 DB 953 VQFTSITVYNNARPQSDITLTFADINNDPKLISQCFKKGKKEATIOF 1001

RESULT 13
 MYT_MOUSE
 ID MYT_MOUSE STANDARD; PRT: 1099 AA.
 AC P70248;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin If.
 GN MYOIF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RC STRAIN-BALB/C; TISSUE-Cochlea;
 RC MEDLINE=97237053; PubMed=9119401;
 RA Crozet F., Amraoui A.E., Blanchard S., Lenoir M., Ripoll C., Vago P.,
 RA Hamel C., Fizesmes C., Levi-Acobas F., Depietis D., Mattei M.-G.,
 RA Weil D., Pujol R., Petit C.;
 RT Cloning of the genes encoding two murine and human cochlear
 RT unconventional type I myosins.";
 RL Genomics 40:332-341(1997).
 CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
 CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
 CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
 CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
 CC SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, SPLEEN, EYE,
 CC BRAIN, LUNG, SMALL INTESTINE, TESTIS AND COCHLEA. BARELY
 CC DETECTABLE IN HEART.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X97650; CAA66251.1; -
 CC HSSP; P08799; 1MND.
 CC MGD; MG1:107711; Myoif.
 CC InterPro: IPR00048; IO.region.
 CC InterPro: IPR001452; SH3.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam; PF00018; SH3; 1.
 CC Pfam; PF00063; myosin_head; 1.
 CC Pfam; PF00612; IO; 1.
 CC PRINTS; PR00193; MYOSINHEAVY.
 CC PRINTS; PR00452; SH3DOMAIN.
 CC ProDom; PD000066; SH3; 1.
 CC ProDom; PD000355; myosin_head; 1.
 CC SMART; SM00015; IO; 1.
 CC SMART; SM00242; MYSC; 1.
 CC SMART; SM00326; SH3; 1.
 CC PROSITE; PS0096; IO; 1.
 CC PROSITE; PS50002; SH3; 1.

QY	26	GVODFVLNENFTSEAFIENLRRRRENIIYIYIPVLVSNVXPDLQIYROMENRG	87
Db	18	GVODAVLLLSQIT-EDQIVNRLLHKRRMDVIFIIYISVLISVNPFGOMYFIEDREIDLXG	76
QY	88	VSEYEVPHLFAVADTVYALRTERDDQAVMISGESGAKTEATKRLQFYAETCPAPER	147
Db	77	AVQEBEPPIYALTDMYEMMLIDENOCVIIISGESGAKTVAAYKIMGYISKSGGGK	136
QY	148	GGAVRRLLQSNVPLEAFGNACTLRNDNSRFGKXNDVOPFGKAPVGHLSLVLKSR	207
Db	137	VQVHKIILQSNPLEAFGNACTVRNNNSRFGKTFEIJQSRGSGPDGKISNLEKSR	196
QY	208	VHQHNGERNFHVYQOLLEGGEETLRRLGLERNPOSYLVKGCQAVSSINXSDMKV	267
Db	197	VVMQNNERNFHIYQOLLEGASQEQOQNLGI-MSQDYUYUYNQSDPYKVEGTDDRDPSE	25
QY	268	MKRLSLVIDTDEVEDLLSIYASVLHGNIHFADEDSMAQVTTENQKLTLLLQVEG	327
Db	256	TLSAMQVMBGRACQOOLYQDLVAGLHGNISF-CEENYVRVSVSLAPAYLLIDS	314
QY	328	TYLREALTFHRIIAK---GELLSPLENGQAARYDRLAAVYSRPTVWLRKINSLA	383
Db	315	GRQELKTRKNDKSKMGGRSESDITVLNQAQARYDRLAGVRLAFDELVEALINAM-	373
QY	384	SKQAESPRMSRTTVJGLDIYGFYVQHNSEFOFCINCNELQOLFIELTLKSEOEYE	443
Db	374	QKQOEYS-----IGVDIYGFELFQKNGEFOFCINVENKQOIFIELTLKQOEYV	427
QY	444	AEBIAEPQYQFNKIKIDVBEKK--GIISILDECL---RGENDTLFELKEDIV	498
Db	428	QEBIRKTPLEIFENNKIVCDLIENKISPREJISVLDVATYMHATGGADQYLQKLOAAV	487
QY	499	KPHPHLTKHLADOKTRKSLDGEFRLHYAGEVYSVTGFELDKNNDLFRNLKETQSS	558
Db	488	GTEHEH-----NSMSAG-FYIINHAGKVSVYSGFCERNRDLVFDLELMQSS	535
QY	559	MNPIMQCF-DKSELSDDKRPETVATQRMSLDQVELLSKEPAYTRICIKRPNDAKOPR	617
Db	536	DODFLMLRPEPKLNTDKGRPSTASKIKIKQANDVSYTLKCTPHYIKICIPNETKPRD	595
QY	618	FDEVLTRHONVKLIGLMENTRVRRAGFAVRKYEALLOYKSLCQETPMWMAGRQDQVAV	677
Db	596	WESRKHQVETLGLRENIRVVRAGFAIRKQPSKLOKXALITPEYTPMRMGDEROGQH	655
QY	678	LVRLHGYKPEEYKMGRTKLEIRPKTLRATEDSLSEVRQSATKIQOAMRFGHROKFLR	737
Db	656	ILRAVNMEDQYQMGSTKYFVKNPESLFLLEMBRERKFDGPARTIQAM-----	704
QY	738	VKRSALCIQSNMGRGLGRKKAKKRKYMAOTIRLRLGFIILHSPRCBNAFELLDHVASF	797
Db	705	-----RRHVAVRY-----	713
QY	798	LNLRLRQRLPRNYLDTSMPPRALKEASELLRELCMKMMWKKYCRSISPWKQOLOKAV	857
Db	714	-----EENRE-----E	719
QY	858	ASBIEFGKADNTPQSVPRLEISTRLGTEISPRVLOSGS--EPQIYAVPVYKIDRYKPR	916
Db	720	ASLILNKKERRRNSINRNFGVNYGLEE-RELEHQFLAKRERYDFQDSVYKYDR-FKP	777

```

RESULT 14
MYSB-DICDI STANDARD; PRT; 1111 AA.
AC P34092;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MYOSIN IB heavy chain.
GN MYOA OR DMID.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=89345628; PubMed-2763220;
RA Jung G., Saxe C.L. II, Kimmel A.R., Hammer J.A. III;
RT "Dictyostellium discoideum contains a gene encoding a myosin I heavy
   chain.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6186-6190(1989).
RL RN 12;
RP SEQUENCE OF 481-490; 656-666 AND 783-798.
RC STRAIN=AX3;
RX MEDLINE=93315475; PubMed-8325874;
RA Jung G., Fukui Y., Martin B., Hammer J.A. III;
RT "Sequence, expression pattern, intracellular localization, and
   targeted disruption of the Dictyostellium myosin I heavy chain
   isoform.";
RL J. Biol. Chem. 268:14981-14990(1993).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
   ACTIVITY THAT IS ACTIVATED BY ACTIN. MYOSIN IB MAY HAVE A ROLE IN
   CHEMOTAXIS AND AGGREGATION. IT COULD SERVE TO STABILIZE AND EVEN
   RETRACT CORTEXICAL STRUCTURES, SUCH AS PSEUDOPODS AND LAMELLOPODS.
CC -1- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY
   AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION JUST BENEATH THE
   PLASMA MEMBRANE IN THE ANTERIOR PSEUDOPOD AT THE LEADING EDGE OF
   THE CELL.
CC -1- DOMAIN: TH.1 BINDS DIRECTLY TO ANIONIC PHOSPHOLIPID MEMBRANES;
   MYOSINS I COULD THEREFORE MOVE ACTIN RELATIVE TO MEMBRANES AND
   VICE VERSA. TH.2 AND SH3 BIND TIGHTLY TO F-ACTIN; THIS TOGETHER
   WITH THE NUCLEOTIDE-SENSITIVE SITE IN THE HEAD, ALLOWS SINGLE
   MOLECULES OF MYOSIN I TO CROSS-LINK ACTIN FILAMENTS.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@sib-ch).
CC -----
DR EMBL, M26037; AAA33229.1; -
DR PIR, A33284; A33284.
DR HSP: P08799; 1MND.
DR DictyDb; DD01047; myoa.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001452; SH3.

```

```

DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00018; SH3_1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00612; IQ; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRINTS: PR00452; SH3DOMAIN.
DR ProDom: PD000066; SH3_1.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00242; MYSC; 1.
DR SMART: SM00226; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
DR Myosin: Actin-binding; ATP-binding; SH3 domain; Multigene family;
KW Hemotaxs; Phosphorylation.
FT DOMAIN 1 694 MYOSIN HEAD-LIKE.
FT DOMAIN 922 1052 TAIL HOMOLOG REGION 1 (TH.1).
FT DOMAIN 1053 1111 GLY/PRO/ALA-RICH (TH.2).
FT NP_BIND 102 109 SH3.
FT NP_BIND 547 627 ATP (POTENTIAL).
FT DOMAIN 951 1015 ACTIN-BINDING.
FT MOD_RES 332 332 ASN-RICH.
SO SEQUENCE 1111 AA; 124313 MW; CD6386F08DC3642P CRC64;

Query Match 31.4%; Score 1692.5; DB 1; Length 1111;
Best Local Similarity 36.4%; Pred. No. 9.9e-100;
Matches 384; Conservative 172; Mismatches 312; Indels 187; Gaps 21;

QY 28 GVQDFVLENTSEAEFIENLRFRRENLIYTYIGVLYSVNRYRLOIYRQHMERYG 87
DB 10 GTDLDVLPK-VSEDEICENLKRKRYMDFITYNIGVLYSVNRYRLOIYRQHMERYG 68
QY 88 VSEFEVPHLFAVDYFRRALRTERRDQAVMISGSGAGTEKRLLOFYAETCAPER 147
DB 69 KHAQEVPHYQALAESYRKMKNQENQCYIIGSEGAGTEAKLIMGVSAISGSTER 128
QY 148 GGAVRDLLOSNPVLEAFNAKTLRDNSSREKGYNDVQDFEGAPVGHISLYLEKSR 207
DB 129 VEYKHVILSNPLLEAFNAKTLRNNSSRFKGYEIOFDKAGDVGKITYNYLLEKSR 188
QY 208 VVHONHGRNFHYVYOLLEGESEETLRGLERNPOSYLYLVGCAKYSINDKDMY 267
DB 189 VVYONHGRNFHYVYOLLEGESEETLRGLERNPOSYLYLVGCAKYSINDKDMY 247
QY 268 MKRALSYDTEDEVEDLSTIVASVHLGNHRADEDSNAQYTTENOLKYLTRLGVGR 327
DB 248 VROAMDTIGLTAQDQSDIIRIVACVHLGINVFIETEDKQNAALYDPNALSLASMLCIDS 307
QY 328 TITREALTHKRIAKG-----ELLSPLEQAAVADALAKAVYKSTFTWLYRKIN 379
DB 308 ATLONALLFRVINTGAGGAGNRRSTYVNPQNEQANGTRDALARTIYDMFSLVEKYN 367
QY 380 RSLASKADESPSRSTVGLLDIYGEVQHNHSPFOFCINYNENKLOOLFETLTKSRQ 439
DB 368 QSLSY-YKSP---YQNVIGILDIIFGEIERKNGFEOFCINYNENKLOOLFETLTKSRQ 422
QY 440 EEEYAEGLAPVQYFNKIKICDLVEEKF-KGIISLDECL--RPGATYDLTLEKLE 495
DB 423 EEEYREGIKWEPIKYEYFNQIVCDLIECKSPGIFSLDDICSLHAGSTQDCKLEKKA 482
QY 496 DYKPRPHF--LTHKLADQKTRSLDGERLLHYGEVYISTVGLDKNDLLFRNLKE 553
DB 483 GIYDGLHMRGM-----GAFALHYAGEVYEEYEGESDKKDTLFDLIE 528
QY 554 TMSQSNPIAACOFDEKSELS-DKRPETVATQKMSLDQVELLRKSEPAYIRICIPNA 612
DB 529 AIOCKMKPFLASIFNEGTSLQKKRPTTAGFKIKITAGELMKLSOCTPHYIKICIPNET 588
QY 613 KQGRFDEVLIRHQVYKILGIMENLYRRAGFAVRRYEAFLQRYKSLCETWPMMA---G 669
DB 589 KKAADMENSVKQVOYIGLLENVVRRAAGFAIRNFDKYLKRYKSLSTWIGIEMWNG 648
QY 670 RPDGVAVLVRLHYRPEEYKMGRTYIFIRPKTLFATEDSELEVRQSLATKIQAAKGP 729
DB 649 DAIEGCKTIFQDMNLEAGOWLGKTKVFLRHEPVEFLLEALDKKDFDCTAKIQAFR-- 706
QY 730 HWROKFLRKRAICIQSWMRGTLGRRAAKKMAQIIRLIRNGFILRHSPCEANFF 789
DB 707 MNRAR----- 711
QY 790 LDHVASFELNLRQLPRNVLDTSWPTPPALREASELLRELCMKNNVWKYCRSISPBMK 849
DB 712 -----K 712
QY 850 QOLOKAAVASELFKKKNYPOSVRLISTRLGTEF--ISPRYQSLSEPTQYAVPY 906
DB 713 HSELEORAOIAHMFKKKERQNSIDRKTSVIDEFENQGOEQMNNKREKERVYFADTV 772
QY 907 VYDGRKGYRPRROLLTPSAVYVEDKKV-----ORDVNLGISVSSLSDSLF 958
DB 773 IKIDRA-KQKYEAVLTDQALYFEKSIKKVVLHVLIRRGLEIGVSIISTLSDVY 831
QY 959 VLVHVRD-----NKQGVVLOSDHVIETLFTALS--ADRVNNININQSGITFAGPG 1011
DB 832 VHLHEHQVLENDKTEILLVVEYFAIGGSLNVQFSDRIN-----YTLKKGEO 883
QY 1012 RKGITDFTSGSE--LLITKAKGHLAVAPRLNS 1043
DB 884 KE--ISFQSEQCPVLVVKGGKGLIGTASGLPS 916

```

```

RESULT 15
MISD.DICDI
ID MISD.DICDI STANDARD; PRI; 1113 AA.
AC P34109;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin ID heavy chain.
GN MYOD OR DMTD.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NC NCB1_Taxid=44689;
RN [1]
RC SEQUENCE FROM N.A., AND SEQUENCE OF 604-610; 733-742 AND 914-928.
RA STRAIN=AX3;
RX MEDLINE=93315475; PubMed=8325874;
RA Jung G., Fukui Y., Martin B., Hammer J.A. III;
RT "Sequence, expression pattern, intracellular localization, and
RT targeted disruption of the Dictyostelium myosin ID heavy chain
RT isoform.";
RL J. Biol. Chem. 268:14981-14990(1993).
CC -I- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY ACTIN. MYOSIN ID MAY HAVE A ROLE IN
CC CHEMOTAXIS AND AGGREGATION; IT COULD SERVE TO STABILIZE AND EVEN
CC RETRACT CORTICAL STRUCTURES, SUCH AS PSEUDOPODS AND LAMELLOPODS.
CC -I- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY
CC AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.
CC -I- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION JUST BENEATH THE
CC PLASMA MEMBRANE IN THE ANTERIOR PSEUDOPOD AT THE LEADING EDGE OF
CC THE CELL.
CC -I- DOMAIN: TH.1 BINDS DIRECTLY TO ANIONIC PHOSPHOLIPID MEMBRANES;
CC MYOSIN I COULD THEREFORE MOVE ACTIN RELATIVE TO MEMBRANES AND
CC VICE VERSA. TH.2 AND SH3 BIND TIGHTLY TO F-ACTIN. THIS TOGETHER
CC WITH THE NUCLEOTIDE-SENSITIVE SITE IN THE HEAD, ALLOWS SINGLE
CC MOLECULES OF MYOSIN I TO CROSS-LINK ACTIN FILAMENTS.
CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



```

Db      1  MRYRASALGSDGVRTMESALTARDVGVODVLENTSEAFITNLRREPLNTY 60
QY      61  IGVPLVSNPFDLQITSRQHMERYGVSEVPPHLEAVADTVRALTERRDQAVMTS 120
Db      61  IGVPLVSNPFDLQITSRQHMERYGVSEVPPHLEAVADTVRALTERRDQAVMTS 120
QY      121  GESGAGTEATKRLDLPYATPCAPERGAVRDLQSNPVEAFNATLRDNDNSRRG 180
Db      121  GESGAGTEATKRLDLPYATPCAPERGAVRDLQSNPVEAFNATLRDNDNSRRG 180
QY      181  KYMDVDFEKGAVGGHISYLEEKSRYVQHNGERNHFYOLLEGGSEETLRRLGLER 240
Db      181  KYMDVDFEKGAVGGHISYLEEKSRYVQHNGERNHFYOLLEGGSEETLRRLGLER 240
QY      241  NPOSYLYLVKGCACAKVSSINDKSDMKVMKALSVIDFTEDEVEDLSTIASVHLGNH 300
Db      241  NPOSYLYLVKGCACAKVSSINDKSDMKVMKALSVIDFTEDEVEDLSTIASVHLGNH 300
QY      301  AAEDESNAGVTENOLKYLRLGVTGTLREALTRKTIKAGEELLSPNLBOAAYARD 360
Db      301  AAEDESNAGVTENOLKYLRLGVTGTLREALTRKTIKAGEELLSPNLBOAAYARD 360
QY      361  ALAKAVYSRTFTWLVRKINSLSKDAESPMSSTVGLDLYGFEVQHNSEQFCIN 420
Db      361  ALAKAVYSRTFTWLVRKINSLSKDAESPMSSTVGLDLYGFEVQHNSEQFCIN 420
QY      421  YCNKRLQQLFIELTILKSEOEYEAEGIAMPEVOYFNKKIICDLVEEKFGIISLDE 480
Db      421  YCNKRLQQLFIELTILKSEOEYEAEGIAMPEVOYFNKKIICDLVEEKFGIISLDE 480
QY      481  RPEADTDLFLEKLEEDVVKPHPHFLTHKLADQKTRSLDGEFRLHYAGEVYVYTG 540
Db      481  RPEADTDLFLEKLEEDVVKPHPHFLTHKLADQKTRSLDGEFRLHYAGEVYVYTG 540
QY      541  DKNNDLFRNLKETTMOSSNMPIMAOCFKSELSDDKRPETVATQFKMSLQVLVLR 600
Db      541  DKNNDLFRNLKETTMOSSNMPIMAOCFKSELSDDKRPETVATQFKMSLQVLVLR 600
QY      601  PAYIRCTKPKDAPKOPGFEDVLIRHOKYVGLMENVLRVAGAVYRKRYEAFQRYK 660
Db      601  PAYIRCTKPKDAPKOPGFEDVLIRHOKYVGLMENVLRVAGAVYRKRYEAFQRYK 660
QY      661  PETWPMWAGRPDGVAVLVNHLGYKPEEYKMGRTKIFIRPKTLFATEDSLLEV 720
Db      661  PETWPMWAGRPDGVAVLVNHLGYKPEEYKMGRTKIFIRPKTLFATEDSLLEV 720
QY      721  KIQAAWGFHWROKFLVKKSAICIOSWMTGLGRRAAKRKMAAOTIRRLINGFLIR 780
Db      721  KIQAAWGFHWROKFLVKKSAICIOSWMTGLGRRAAKRKMAAOTIRRLINGFLIR 780
QY      781  PRCEENAFIDHVRASFLNLRQLPNNVLDTSMPPTPALREASELRLCKNNVMYK 840
Db      781  PRCEENAFIDHVRASFLNLRQLPNNVLDTSMPPTPALREASELRLCKNNVMYK 840
QY      841  CRISPEMKOOLQOKAVASEIFKGGKNDNPQSVPRLESTRIGTEESIPRYLOSL 900
Db      841  CRISPEMKOOLQOKAVASEIFKGGKNDNPQSVPRLESTRIGTEESIPRYLOSL 900
QY      901  QYAVPVYKDYKGYKPPRQLLTPSAVYVEDAKYKQRIIDYANLIGISSLSLFLV 960
Db      901  QYAVPVYKDYKGYKPPRQLLTPSAVYVEDAKYKQRIIDYANLIGISSLSLFLV 960
QY      961  HVQEDNRKQGVLOSADHVIETLTALTSADRVNININOGSITFAGGGRGIIIDFT 1020
Db      961  HVQEDNRKQGVLOSADHVIETLTALTSADRVNININOGSITFAGGGRGIIIDFT 1020
QY      1021  GSELLITKAKNGHLAVVAPRLNSR 1044
Db      1021  GSELLITKAKNGHLAVVAPRLNSR 1044

```

RESULT 2
063355

```

ID      063355  PRELIMINARY:  PRT:  1028 AA.
AC      063355;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Myosin I heavy chain.
GN      MYR 2.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SPRAGUE-DAWLEY;
RA      Ruppert C., Godel J., Reinhard J., Baeher M.;
RT      "Myr-2 a novel class-I myosin identified in rat brain.";
RL      Submitted (Aug-1993) to the EMBL/GenBank/DBJ databases.
DR      EMBL; X74800; CAA52807.1; -.
DR      HSSP; P08799; 1MND.
DR      InterPro: IPR000048; IQ_region.
DR      InterPro: IPR001609; myosin_head.
DR      Pfam; PF00612; IQ_3.
DR      Pfam; PF00063; myosin_head; 1.
DR      PRINTS; PR00193; MYOSINHEAVY.
DR      ProDom; PD000355; myosin_head; 1.
DR      SMART; SM00015; IQ; 2.
DR      SMART; SM00242; MYSC; 1.
SQ      SEQUENCE 1028 AA; 118090 MW; 10E3B1AFD5999CB5 CRC64;

Query Match      97.5%; Score 5261; DB 11; Length 1028;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1015; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY      17  MESALTARDVGVODVLENTSEAFITNLRREPLNTYTGPLYVSNPFDLQI 76
Db      17  MESALTARDVGVODVLENTSEAFITNLRREPLNTYTGPLYVSNPFDLQI 76
QY      1  MESALTARDVGVODVLENTSEAFITNLRREPLNTYTGPLYVSNPFDLQI 60
Db      1  MESALTARDVGVODVLENTSEAFITNLRREPLNTYTGPLYVSNPFDLQI 60
QY      77  YSRQHMERYGVSEVPPHLEAVADTVRALTERRDQAVMISGESGAGTEATKRLQ 136
Db      77  YSRQHMERYGVSEVPPHLEAVADTVRALTERRDQAVMISGESGAGTEATKRLQ 136
QY      61  IGVPLVSNPFDLQITSRQHMERYGVSEVPPHLEAVADTVRALTERRDQAVMIS 120
Db      61  IGVPLVSNPFDLQITSRQHMERYGVSEVPPHLEAVADTVRALTERRDQAVMIS 120
QY      137  FYAETCPAPERGAARDRLQSNPVEAFNATLRDNDNSRFGKYMVOVDFGAPYGG 196
Db      137  FYAETCPAPERGAARDRLQSNPVEAFNATLRDNDNSRFGKYMVOVDFGAPYGG 196
QY      121  FYAETCPAPERGAARDRLQSNPVEAFNATLRDNDNSRFGKYMVOVDFGAPYGG 180
Db      121  FYAETCPAPERGAARDRLQSNPVEAFNATLRDNDNSRFGKYMVOVDFGAPYGG 180
QY      197  HILSTYLEKSRVYVQHNGERNHFYOLLEGGSEETLRRLGLERNPQSYLYLVK 256
Db      197  HILSTYLEKSRVYVQHNGERNHFYOLLEGGSEETLRRLGLERNPQSYLYLVK 256
QY      181  HILSTYLEKSRVYVQHNGERNHFYOLLEGGSEETLRRLGLERNPQSYLYLVK 240
Db      181  HILSTYLEKSRVYVQHNGERNHFYOLLEGGSEETLRRLGLERNPQSYLYLVK 240
QY      257  SSINDKSDMKVMKALSVIDFTEDEVEDLSTIASVHLGNHFAADESNAGVTEN 316
Db      257  SSINDKSDMKVMKALSVIDFTEDEVEDLSTIASVHLGNHFAADESNAGVTEN 316
QY      241  SSINDKSDMKVMKALSVIDFTEDEVEDLSTIASVHLGNHFAADESNAGVTEN 300
Db      241  SSINDKSDMKVMKALSVIDFTEDEVEDLSTIASVHLGNHFAADESNAGVTEN 300
QY      317  KYTLRLGVEGTLREALTRKTIKAGEELLSPNLBOAAYARDALAAVSRFTWLV 376
Db      317  KYTLRLGVEGTLREALTRKTIKAGEELLSPNLBOAAYARDALAAVSRFTWLV 376
QY      301  KYTLRLGVEGTLREALTRKTIKAGEELLSPNLBOAAYARDALAAVSRFTWLV 360
Db      301  KYTLRLGVEGTLREALTRKTIKAGEELLSPNLBOAAYARDALAAVSRFTWLV 360
QY      377  KINRSLASKDAESPMSSTVGLDLYGFEVQHNSEQFCINYNCKNLQDLFELTL 436
Db      377  KINRSLASKDAESPMSSTVGLDLYGFEVQHNSEQFCINYNCKNLQDLFELTL 436
QY      361  KINRSLASKDAESPMSSTVGLDLYGFEVQHNSEQFCINYNCKNLQDLFELTL 420
Db      361  KINRSLASKDAESPMSSTVGLDLYGFEVQHNSEQFCINYNCKNLQDLFELTL 420
QY      437  SEOEYEAEGIAMPEVOYFNKKIICDLVEEKFGIISLDEBCLPFGATDLFLEK 496
Db      437  SEOEYEAEGIAMPEVOYFNKKIICDLVEEKFGIISLDEBCLPFGATDLFLEK 496
QY      421  SEOEYEAEGIAMPEVOYFNKKIICDLVEEKFGIISLDEBCLPFGATDLFLEK 480
Db      421  SEOEYEAEGIAMPEVOYFNKKIICDLVEEKFGIISLDEBCLPFGATDLFLEK 480
QY      497  TYKPRPHLTHKLAQKTRKSLDGEFRLHYAGEVTVSYTGFLDKNNDLFRNLKET 556
Db      497  TYKPRPHLTHKLAQKTRKSLDGEFRLHYAGEVTVSYTGFLDKNNDLFRNLKET 556
QY      481  TYKPRPHLTHKLAQKTRKSLDGEFRLHYAGEVTVSYTGFLDKNNDLFRNLKET 540
Db      481  TYKPRPHLTHKLAQKTRKSLDGEFRLHYAGEVTVSYTGFLDKNNDLFRNLKET 540
QY      557  SSNMPIMAOCFKSELSDDKRPETVATQFKMSLQVLVLRKSEKAPYTRCIKPD 616
Db      557  SSNMPIMAOCFKSELSDDKRPETVATQFKMSLQVLVLRKSEKAPYTRCIKPD 616
QY      541  SSNMPIMAOCFKSELSDDKRPETVATQFKMSLQVLVLRKSEKAPYTRCIKPD 600
Db      541  SSNMPIMAOCFKSELSDDKRPETVATQFKMSLQVLVLRKSEKAPYTRCIKPD 600

```


DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Myosin IB.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=94294379; PubMed=8022785;
 RA Reitzes O., Barylko B., Li C., Sudhof T.C., Albanist J.P.;
 RT Domain structure of a mammalian myosin I-beta.
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6349-6353(1994).
 DR EMBL: Z22852; CA80476.1; -.
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR000048; IO_region.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 2.
 DR SMART: SM00242; MYSC; 1.
 SO SEQUENCE 1028 AA; 118241 MW; 95E7E806CC6B2CD CRC64;

Query Match 94.7%; Score 5111; DB 6; Length 1028;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 980; Conservative 27; Mismatches 21; Indels 0; Gaps 0;

QY 17 MESALRARDRVGVQDVLLENFTSEAFTEMLRRRREMLITYIGPVLYVNPYDLOI 76
 DB 1 MESALTCRDRVGVQDVLLENFTSEAFTEMLRRRREMLITYIGPVLYVNPYDLOI 60
 QY 77 YSHQHMERKGVGFYEVPPHFAVDVYRALTERERDQVAMISGSGAGKTEATRLIQ 136
 DB 61 YSHQHMERKGVGFYEVPPHFAVDVYRALTERERDQVAMISGSGAGKTEATRLIQ 120
 QY 137 FYAETCPAPERGAVDRILQSNPVLEAFGNATKLNDSRRGKTYMDVQDFKGA PVG 196
 DB 121 FYAETCPAPERGAVDRILQSNPVLEAFGNATKLNDSRRGKTYMDVQDFKGA PVG 180
 QY 197 HILSTLEKSRVYHONHGERNFHFYQLLEGGEETRLGLERNOSLYLVKGCACV 256
 DB 181 HILSTLEKSRVYHONHGERNFHFYQLLEGGEETRLGLERNOSLYLVKGCACV 240
 QY 257 SSINDKSDMKVNRKALSVIDFTEDEVEDLSTIYASVLAHNTIFADEDSMAQVTEENQ 316
 DB 241 SSINDKSDMKVNRKALSVIDFTEDEVEDLSTIYASVLAHNTIFADEDSMAQVTEENQ 300
 QY 317 KYTRLRGVGTTLREALTHRTKIANGEBELSPLEQAAVARDALAKAVYSTRFTWLV 376
 DB 301 KYTRLRGVGTTLREALTHRTKIANGEBELSPLEQAAVARDALAKAVYSTRFTWLV 360
 QY 377 KIRNSLASKDAESPMSRSTTVGLILTYGFEVQHNHSPFCINYNENKLOOFIELTLK 436
 DB 361 KIRNSLASKDAESPMSRSTTVGLILTYGFEVQHNHSPFCINYNENKLOOFIELTLK 420
 QY 437 SEOEYEAEGIAIEPVOYFNKLTICDVEKEFGIISILDECLRGEATDITFLEKLD 496
 DB 421 SEOEYEAEGIAIEPVOYFNKLTICDVEKEFGIISILDECLRGEATDITFLEKLD 480
 QY 497 TVKPHFLTHKALADOKTRKSLDRGEFLIHAAGEVTVSYTGFLKNNDLRLNKEIYC 556
 DB 481 TVKPHFLTHKALADOKTRKSLDRGEFLIHAAGEVTVSYTGFLKNNDLRLNKEIYC 540
 QY 557 SSNNPIMAOOFKSESDKKRPETVTOFMSLSLOVELIRSEPAVYICIPNDKAKOG 616
 DB 541 SSNNPIMAOOFKSESDKKRPETVTOFMSLSLOVELIRSEPAVYICIPNDKAKOG 600
 QY 617 RFDEVLIHQVYKYLGIEMNLVRRAGFAVRRKYEAFLOQRYKSLCPETPMWAGRPDGYA 676

DB 601 RFDEVLIHQVYKYLGIEMNLVRRAGFAVRRKYEAFLOQRYKSLCPETPMWAGRPDGYA 660
 QY 677 VLVRLHGYKPEEYKMGRTKIFIRPKTLFATEDSLEVRQSLATKIOAMRGFWHROKFL 736
 DB 661 VLVRLHGYKPEEYKMGRTKIFIRPKTLFATEDSLEVRQSLATKIOAMRGFWHROKFL 720
 QY 737 RVKRSALICISQWMTGGRRAAKRKMAQITRLNGFLIRHSRCPENAFIDYHRAS 796
 DB 721 RVKRSALICISQWMTGGRRAAKRKMAQITRLNGFLIRHSRCPENAFIDYHRAS 780
 QY 797 FLNLRQLPNNVLDISWPTPPALREASELLRELCNMWYKCRSISPENKQOOLQOKA 856
 DB 781 FLNLRQLPNNVLDISWPTPPALREASELLRELCNMWYKCRSISPENKQOOLQOKA 840
 QY 857 VASEIFKCKNDYQSVPRLEFISITRLGTEISPRVLOSGLSEPIQIYAVPVYKIDRKGYK 916
 DB 841 VASEIFKCKNDYQSVPRLEFISITRLGTEISPRVLOSGLSEPIQIYAVPVYKIDRKGYK 900
 QY 917 RPRQLLPNSAVVYEDAKVKORIDYANLTGISVSLSDSLFVLYHOREDNKOKGDPVLO 976
 DB 901 RSRQLLPNSAVVYEDAKVKORIDYANLTGISVSLSDSLFVLYHOREDNKOKGDPVLO 960
 QY 977 SDHVIELTKTALSADRVNNININGSLTFAGGGRGDIIDFTSGSELLTKRANGHLAV 1036
 DB 961 SDHVIELTKTALSADRVNNININGSLTFAGGGRGDIIDFTSGSELLTKRANGHLAV 1020
 QY 1037 VAPRLNSR 1044
 DB 1021 VAPRLNSR 1028

RESULT 5

ID 092002 PRELIMINARY; PRT: 1028 AA.
 AC 092002;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Myosin I beta.
 GN Aml beta.
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=95083594; PubMed=7991542;
 RA Metcalf A.B., Chelliah Y., Hudspeth A.J.;
 RT "Molecular cloning of a myosin I beta isozyyme that may mediate
 adaptation by hair cells of the bullfrog's internal ear."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11821-11825(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Solc C.F., Deflier B.H., Duyk G.M., Corey D.P.;
 RL Aud. Neurosci. 1:63-75(1994).
 DR EMBL: U14549; AA57192.1; -.
 DR EMBL: U14382; AA65091.1; -.
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR000048; IO_region.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 2.
 DR SMART: SM00242; MYSC; 1.
 SO SEQUENCE 1028 AA; 118830 MW; D15F6F99B86A90AE CRC64;

Query Match 81.0%; Score 4371; DB 13; Length 1028;
 Best Local Similarity 79.2%; Pred. No. 1; 2e-292;

Matches	814;	Conservative	112;	Mismatches	102;	Indels	0;	Gaps	0;
OY	17	MESALTARDVGVODVLENTSEAFIENLRRERENITYITGPVSVNPRDLOI	76						
DB	1	MESALTARDVGVODVLENTSEAFIENLRRERENITYITGPVSVNPRDLOI	76						
OY	77	YSHQHMERGVSEFVEVPHLFAVADTVYRALTEREDQVMSISGSGAGTEATRLLO	136						
DB	61	YSHQHMERGVSEFVEVPHLFAVADTVYRALTEREDQVMSISGSGAGTEATRLLO	136						
OY	137	FVAFETCPAREGAVADRLQSNPVLEAFGNKTLNDSSRFKGYMDVQDFKAPVGG	196						
DB	121	FVAFETCPAREGAVADRLQSNPVLEAFGNKTLNDSSRFKGYMDVQDFKAPVGG	196						
OY	197	HILSYLLEKSRVYVHONHGERNFHFVQLEEGEERTLRGLGERNQSLTYLVKGCACV	256						
DB	181	HILSYLLEKSRVYVHONHGERNFHFVQLEEGEERTLRGLGERNQSLTYLVKGCACV	256						
OY	257	SSINDKSDMKVAKALSVIDFTEDEVEDLLSIVASVHLGNHFADEDSMAQVTTENOL	316						
DB	241	SSINDKSDMKVAKALSVIDFTEDEVEDLLSIVASVHLGNHFADEDSMAQVTTENOL	316						
OY	317	KYTLRLGEGTTLREALTHRKTIAGEELLSPLENOQAAYADALAKAVYKRTFWLVR	376						
DB	301	KYTLRLGEGTTLREALTHRKTIAGEELLSPLENOQAAYADALAKAVYKRTFWLVR	376						
OY	377	KINRSIASKDAESPSMRSTTVGLDIDYGFVEVQHNSEFOFCINCEKLOOLELTLK	436						
DB	361	KINRSIASKDAESPSMRSTTVGLDIDYGFVEVQHNSEFOFCINCEKLOOLELTLK	436						
OY	437	SEDEYEAAGIAMEPVQYFNKIIICDLVEEKGIIISILDEBCLRGATDLPFEKLED	496						
DB	421	SEDEYEAAGIAMEPVQYFNKIIICDLVEEKGIIISILDEBCLRGATDLPFEKLED	496						
OY	497	TYVPHPHFTHKLAODKTRKSLDRGEFLLHAGETVSVTFIDKNNDLRLNLEKTEVC	556						
DB	481	TYVPHPHFTHKLAODKTRKSLDRGEFLLHAGETVSVTFIDKNNDLRLNLEKTEVC	556						
OY	557	SSMNPIMACFDKSELSDKRRETVATQFKMSLDQVELLRSKEPAPIICIRPNDAKOPG	616						
DB	541	SSMNPIMACFDKSELSDKRRETVATQFKMSLDQVELLRSKEPAPIICIRPNDAKOPG	616						
OY	617	REDEVILRHQVYTLGLIENVRVRAGFAFRKYEIQLORYKSLCPTWPMMDGRAMDVA	676						
DB	601	REDEVILRHQVYTLGLIENVRVRAGFAFRKYEIQLORYKSLCPTWPMMDGRAMDVA	676						
OY	677	VIVRHLGYKPEEKMGRTKIFIRFPTLTPATEDSLEVRROSLATKIOAMRGFHRKOKFL	736						
DB	661	VIVRHLGYKPEEKMGRTKIFIRFPTLTPATEDSLEVRROSLATKIOAMRGFHRKOKFL	736						
OY	737	RVKRSALICIOSWMRGTLGRKKAARMAQOTIRLIRGFIIRHSPRCENAFELDHVAS	796						
DB	721	RVKRSALICIOSWMRGTLGRKKAARMAQOTIRLIRGFIIRHSPRCENAFELDHVAS	796						
OY	797	FILNLBROLBRNVLDTSWPTPPALREASSELLRELCKMVMKVKYCRSISPEKKOOLQKA	856						
DB	781	FILNLBROLBRNVLDTSWPTPPALREASSELLRELCKMVMKVKYCRSISPEKKOOLQKA	856						
OY	857	VASEIFKGGKNDYPOSVPRFLISTRLGTEIEISPRVLOSIGSEPIQYAPVNVYDKRGYK	916						
DB	841	VASEIFKGGKNDYPOSVPRFLISTRLGTEIEISPRVLOSIGSEPIQYAPVNVYDKRGYK	916						
OY	917	RRROLLPFSAAVYIVEDAKVKORIDYANLTGISVSSLSLFLVHVGREDNKOKDVALQ	976						
DB	901	RRROLLPFSAAVYIVEDAKVKORIDYANLTGISVSSLSLFLVHVGREDNKOKDVALQ	976						
OY	977	SDHVEITLKTALSADRVNINININGSTFAGCPRGDITDTSSELLITRAKNGHLY	1036						
DB	961	SDHVEITLKTALSADRVNINININGSTFAGCPRGDITDTSSELLITRAKNGHLY	1036						
OY	1037	VAPRLNSR 1044							
DB	1021	VAPRLNSR 1028							

RESULT	6
08T08	PRELIMINARY; PRT: 1035 AA.
ID	08T08
AC	08T08
DT	01-JUN-2002 (Tremblrel. 21, Created)
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE	GH04201P.
GN	MY061F.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCH_TaxId=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-BERKELEY;
RA	Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,
RA	Champe M., Chavez C., Dorsett V., Fafian D., Flise E., George R.,
RA	Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA	Nunco J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA	Yu C., Lewis S.E., Rubin G.M., Celinker S.,
RL	Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
DR	EMBL: AY069044; AAL39189.1;
SO	SEQUENCE 1035 AA; 119023 MW; 6EE2C81E6456E63C CRC64;
Query Match	48.7%; Score 2629; DB 5; Length 1035;
Best Local Similarity	51.1%; Pred. No. 1.8e-172;
Matches	533; Conservative 180; Mismatches 306; Indels 24; Gaps 8;
OY	8
DB	1
OY	68
DB	61
OY	128
DB	121
OY	188
DB	181
OY	248
DB	241
OY	308
DB	300
OY	368
DB	360
OY	428
DB	417
OY	488
DB	477
OY	548
DB	537
OY	608

```

Db      597 KPNDDQJANVFNDYLVLLHQVKYLGIMENLVRBAGFAKRYTELELERSKSTWPNY 656
      668 --AGRFQDGVAVLVRLHLYGKPEEYKMGRTKIFIRPEKTLFATEDSELEVAROSIATKIOAA 725
      657 KRGPGKAGVQOVLVDLGDWDEKRYNGEKLEFIRNPRILFDIEDAVQEKHEIAIQAHA 716
      726 WNGFHWKROFLKVKNSAICISGMWGTGRRKAAKRWAAQTIRLIRGFIIRHSPRDE 785
      717 MKGLMORRRKYLKRAQVILIMOSYCRKLAQAARKREAAADKIRAFIKGFIIRNDAPNGF 776
      786 NAFELDVAASFLNLRQLPNNVLDTSMPTRPPLAREASELLRELCMKNNWVKCRSIS 845
      777 NEEFINAKRMMLRLAKELPTKVLDDKSPHAPGCEESGILHRLHRLHRLKRLIT 836
      846 PEMQOAOQKAAVASELFGKKNYPOSVPLFISTRIGTEI--SPRYLQISGSEPIQY 902
      837 POKRQFELKVLAEKVEKGNKNVYASVSTWFOEDRIPKEHIOVNDVFASTFGSEQLKY 896
      903 AYPVAVYDKRGYKPPRROLLTPSAVYVEDAKVQ-----RIDYALITGISVSLSD 955
      897 QSECTFEDRHGKSRDRFILLSKALYVLDGKTYQKHLPLDKIDE-----TLTNND 950
      956 SLFVLHVOREDNKOKGDVYLOSDDVLETTLTALSDRVNNINI--NQGSIIFAGSGPRDG 1014
      951 DLMVIRIPLDLKDKGDLILIPRIEFSTYIIDYGTASIVSYVDRNLHNHYVKGKG 1010
      1015 IIDFTSGSELLITKAKNGLAVY 1037
      1011 VIDITGAEFGVVRDK--GHLVII 1032

```

RESULT 7

```

ID      063357      PRELIMINARY;      PRT;      1006 AA.
AC      063357;
DT      01-NOV-1996 (Tremblrel. 01, Created)
DT      01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT      01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE      Myosin I.
GN      Myr4.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SPRAGUE-DAWLEY;
RX      MEDLINE=94308268; PubMed=8034741;
RA      Baehler M., Kroschewski R., Stoeffler H.E., Behrmann T.;
      "Rat myr4 defines a novel subclass of myosin I: Identification,
      RT      distribution, localization, and mapping of calmodulin-binding sites
      with differential calcium sensitivity.";
      RT      J. Cell Biol. 126:375-389(1994).
      RL      EMBL; X71997; CAA50871.1; -.
      DR      HSSP; P08799; 1MND.
      DR      InterPro: IPR000048; IO region.
      DR      Pfam; PF00612; IQ; 2.
      DR      PRINTS; PR00193; MYOSINHEADY.
      DR      ProDom; PD000355; myosin_head; 1.
      DR      SMART; SM00015; IQ; 1.
      DR      SMART; SM00242; MYSC; 1.
      SQ      SEQUENCE 1006 AA; 116137 MW; 9AA0B626A0FDA42A CRC64;

```

Query Match 34.5%; Score 1865; DB 11; Length 1006;
 Best Local Similarity 41.7%; Pred. No. 8.5e-120;
 Matches 426; Conservative 158; Mismatches 301; Indels 136; Gaps 26;

```

QY      28 GVQDFVLENTSEAFIEULRRRENLIYTYIGVLYSVNRYRLOQIYSRQHMERYG 87
      10 GKADFVLMDP--VSMPEFMANLRLRFKGRITYPEIGEVVSVNDYKVLINITYGRDITQYKG 68

```

```

QY      88 VSEYEVPHLEFAVADTYVYRALTRERDQAVWISGSEAGKTEATKRLQFAETCAPER 147
      69 RELPERLPAIILDAVYKAKMRKRSKPTCLINISESAGKTEATKRLQFAETCAPER 128
      148 G--GAVRDLQSNPVLAEFANAKTLRNDNSRGKTYMDQFDEKAPVGHILSYLLEK 205
      129 AEIERVKMMLKSCVLEAFPMNATKNNNDNSRGKTYMDINFEKGPISGIGHINNYLLEK 188
      206 SRVYHONHGERNFVYQLEEGEETRLRLGERNPQSYLYLKGCAVSSINDKSDW 265
      189 SRVYQOGESESPHSFYQLQGGSEOMHSHTLOKSSSYNYLVGGAQLK--SSINDAAEF 247
      266 KVMKALSIDFTDEVEDLSTVASYLHGINHFAADES---NAQVTENQKLYLR 321
      248 KVLADAMKVIQKPEEIDTYKILAIHLGNLFYDGDPTLENGK-----VSVIAE 302
      322 LLGVEGTLRLRALTRKLIKAGEELLSPLEQ--AAVARALAKAVYSRFTWLVRKINR 380
      303 LLSTADAVKALYLR--TVATGRDIDKQHTGEQASVGRDAFAKAIYERLECWIVTRIND 361
      381 SLASDAESPSRSTYVGLDIYGFVFOHNSFEQICINCKEQLQOLFTELTKSEOE 440
      362 IIEVKNYDTVHGKNTVIGLDIYGFELDENNSEQICINCKEQLQOLFQVLKQEOE 421
      441 EYEAEGIMAEVQYFNKIIDLVEEKEKGIISILDECLRGEATDLTELEKLEDTVKP 500
      422 EYQREGIPWKHIDFNNQIIVDVEQOKGIALLDQAMNVGAVTGMFLEALNSKLGK 481
      501 HPHELTNRKLDQKTRKSLDGEFRLIYAGVYTVSVGLDKNDLDFRNLKEMGSSM 560
      482 HGHESSRKTCASDKLIEFDR--DFRIIRYAGVYTVSVGLDKNDLDFRNLKEMGSSM 540
      561 PIMACQEPKSELS---DKRPETATOPKMSILOVELTRSEPAVIRICIPNDAKOPGR 617
      541 PVLKMMPEGKLISTEVTKRPPLATLTKNSMIALVNDLASKETPYRYCIKPNOKKSPQI 600
      618 FDEVLIHQVYGLIMENLVRBAGFAVRRRYEAFLOKYSCLCEPWWMAAGPQDGVAV 677
      601 FDDERCRRQVEYGLLENVVRBAGFAVRRRYEAFLOKYSCLCEPWWMAAGPQDGVAV 659
      678 --LVNHLGKPEEYKMGRTKIFIRPEKTLFATEDSELEVAROSIATKIOAMRGHWKQK 735
      660 KKLIRRCFQ--DVAAYGKTKIFIRPRLFTLE--ELRAQMLV----- 699
      736 LKVRSAICISGMWGTGRRKAAKRWAAQTIRLIRGFIIRHSPRCPENAFELDHVRA 795
      700 ----RVVLFQKVMRGITLARMR--YKRTAALITIRYRYKVK----- 737
      796 SFLLNLRQOL--PRNVLD-----TSMPTRPPLAREASELLRELCMKNNWVKCRSISPEWK 849
      738 SYIHVARRFHGVKNMRDYGKHKWPTPPKYLRRFEELQSIIFNRASQILIKTIIPASDL 797
      850 QOQOQKAAVASELFGKKNYPOSVPLFISTRIGTEIISPVLQISGSEPIQYAVPV-- 907
      798 PQVRAKVAAMELKQORADL--GLQRAWEGYTLASKPPTP---OTSGL-----FVPVANE 847
      908 ----KY-----DKRGYKPPRROLLTPSAVYVEDAKV 936
      848 LKRRKKNMNVLESCHVRKVRNRSKVEDAIFVTDBHLTKMP-----TKQVYK 895
      937 KORIDYANLIGISVSLSDSLFVLHVOREDNKOKGDVYLOSDDVLETTLTALSDRVNN 996
      896 KMTIDPLNLTGLSVNGRDOJLVFHT--KDNK-----DLVCLFSKOPTHESTRIGE 944
      997 I 997
      945 I 945

```

RESULT 8

Q27328 PRELIMINARY; PRT; 1017 AA.

AC Q27328:
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE HDM-5 protein.
 GN MCE1A.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Pelodertinae; Caenorhabditis.
 NC NCBI_TaxID=6239;
 RN [1]
 RA SEQUENCE FROM N.A.
 RL Lloyd C.;
 RN Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RP Cope M.J.T.V., Kendrick-Jones J.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z35603; CA84673.1; -;
 DR EMBL: X75564; CA53244.1; -;
 DR HSP: P08799; 1MND.
 DR InterPro: IPR000048; IQ_region.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 KM myosin.
 SQ SEQUENCE 1017 AA; 116557 MW; 461FF63A2C955ED8 CRC64;

Query Match 32.5%; Score 1756; DB 5; Length 1017;
 Best Local Similarity 39.1%; Pred. No. 2.9e-112;
 Matches 405; Conservative 179; Mismatches 329; Indels 124; Gaps 24;

QY 28 GVDFVLLNFTSEAFENLRFRRENLYTYIGPVLYVNPYRDLOYSRQHMERG 87
 DB 12 GVEDVLLSTIDLSV-VQNLQRFQGRITTYIGVLYVANNPYRGLGIVKSVVDYKG 70
 QY 88 VSEYVPPHFAVDYVYRALRTERDDAVMISGSGAGTATKRLQYAEFCAPER 147
 DB 71 REIYERAVHFAIDAAFRSMKRFGRSCIVISGSGAGTETSKITMKYLAITNVROQ 130
 QY 148 G-GAVRDRLLDSNPVLEAFGNATKLRNDSSRFQGVDFDKGAPVGGHIIISYLEK 205
 DB 131 GEIEYKVVNLKSCILFAFCATKTRNDSSRFQGVDFDKGAPVGGHIIISYLEK 190
 QY 206 SRVYHONHGERNFVYQLEGGEEETLRRLGLERNPOSYLYLVKQCAKVSINDKSDW 265
 DB 191 SRVYRQGEGRNFVYQLVNGDGLRQGLKDKAKQYVFLNQGSHKVASINDKSDW 250
 QY 266 KVMKRAL-SVIDFTEDEVEDLSTVASYLHGINHPAEDSN--AOYTENQIKYLR 322
 DB 251 AEVQALALSIHTFPKODVESMWSYAGLIHGNVRFIDGENSSGAVHAEKALQANARC 310
 QY 323 LGVEGTTLREALTRHKITKAGEEELSLPLEQAAYRPAALAKAYSRFTLVKIRSL 382
 DB 311 LNVTPDELAKLSGVVAHGDYKKAQVNAAYTRDALKALERLESMMVSKVNEAI 370
 QY 383 ASKDAEPMRSTVYGLDLYGEVFOHNSFEQFCINCYNEKLOQFETELTLESEOEY 442
 DB 371 SVON--SSRYKSHVIGVLDYGFELFTGNSFEDLCINCYNEKLOQFETELVLAQOEY 428
 QY 443 EAEGTAMPEVOYFNKKITCDLYEKEFKGIISLDECELRPEADLFTLELQTVKPH 502
 DB 429 EREGIKVWKIEYFNKKVYICDLYEIPRGILSLDEACSIGVNDKVFGLGLDKLKSKH 488
 QY 503 HFLTHKLADOKTRKSLDGEFFRLHYAGEVYVYTGFDKNDLFRLLKTCWSSMPI 562
 DB 489 HTSKNTL--KSDKSMGEFEFKITHYAGDYVYVWGFDKDKDILLFQDLKLLHSHKRL 546
 QY 563 MAQCF--DKSELSDKKRPETVAFQFKMSLLQVLEILSKPEPAYIRCIKRPDAKOPGRFD 619

DB 547 VKSLFEDGSKMAEVRRPPTAGFLFKNSMSELVQLAQKEPHYICIPKNEKSNVVD 606
 QY 620 EVLIHQVYILGIMELNRYRRAGFAVRKRYEAFLOWIKSLCEPTWMAGRP-----QD 673
 DB 607 LEREHQVRYLGLLENNVRRAGFAHMPYDRFVNVKILICASTWP-----NPRRGQQLD 662
 QY 674 GVALVRLHLYGKREEKMGRTKIFRPPTLFTEDSLVRQSLATKIQAAWGRPFHMQ 733
 DB 663 SCMOILESGL-AQDVOGRTKIFRSFOTVRLEELRTEQLPNVTFIQKAVRGQOHE 721
 QY 734 KFLVRSALICIOSMWRGTGRRRAKRWMAOTIRRLRGFTLRSLSPRCPENAFELDHV 793
 DB 722 RY-----RMLAVRKITIGAYRRYKLSYIWQ----- 747
 QY 794 RASFLNLRQLPRNVLDTS---WPTPPALREASSELLRELCAKMKWKYCRSIS---P 846
 DB 748 ---LINAERDV--RRMRDLGKSIKRPAPLVLAQFVSRVYMHQ---WRAATILRMP 799
 QY 847 EMKQLOOKRAVASEIFKSKKN-----YQSVPRFLTSTRGTEISPRVLOSIG 896
 DB 800 HLRASLPKATAFEVLANNKEMNGYRRMRGDIYSOELELPTVSTYHDGIALQK-- 857
 QY 897 SEP---IOYAVPVKXDRGKXPRPOLLLTPSAVIVEDAK---VKORIDYANLTGISV 950
 DB 858 SHPFGKVLFTSYQKFN-KENKSLVLYTDRFVAKLENKKFLKEPIPIQISIRISV 916
 QY 951 SLSLDSLEVLHVRED-----NKGQDVLOSQHYIE 982
 DB 917 CAESNGFLFVHGDNDIVGCANKTKNERGVEMIGITLALHYDKITRRSPVLIOS-AVVC 975
 QY 983 TL---TKTALSADRVNN 996
 DB 976 TLGGKTRITVFDENN 992

RESULT 9

ID 090748 PRELIMINARY; PRT; 1099 AA.

AC 090748:
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Brush border myosin IB.
 GN CBMB1B.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 GN Gallus.
 NC NCBI_TaxID=9031;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA TISSUE=INTESTINE;
 RA Knight A.E., Kendrick-Jones J.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL: X70400; CA49850.1; -;
 DR HSP: P08799; 1MND.
 DR InterPro: IPR000048; IQ_region.
 DR InterPro: IPR001609; myosin_head.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR ProDom: PD000355; myosin_head; 1.
 DR ProDom: PD000015; myosin_head; 1.
 DR SMART: SM00242; MYSC; 1.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50002; SH3; 1.

OY		328	TTLRALHLHKKIIA-----KGEELSLPLNLEQAAYARODALAKAVSRTPMLVLRKINRSIA	383
Dd		317	DLREKLTSRKQDSSKWGKSSESHTVLVEQAACYTRDALAKLARHVFDELVSINKAY-	375
OY		384	SKDAESPMSRSTTVGLGLDIYGFEVFNANSFEOFCINCNEKLOOLFETLTKSBOEYE	443
Dd		376	EKHDEHYN-----IGVDIYGFELFGNGRPFQPCINFYNENKLOIFIELTILKAQEELY	429
OY		444	AEGTAMEVOYFNNKIICDDLVEEFK--GIISIDECL--RPGEATDLFFLEKLEDTV	498
Dd		430	QEGIRMPPIEFXNNKIYCDLIENKNPREGSIISLDVOCATMHAVEGAQDTLLKIQMOI	489
OY		499	KPHHFELHKLADOKTRKSLDRGEBRLHYAGELYTSYTGFLDKANNDLFRLKETMGSS	558
Dd		490	GSHHEF-----NSMNQG-FIHNYHGKYSYDMDGFCERRRDVLFMDLILEMOSS	537
OY		559	MNPIMAOCFDKSELSDKK-RPEVTATQPKMSILOVELILRSKEPAVICIKPNDAKOPR	617
Dd		538	ELPFIKSLFPENLQADKGRRPTTAGSKIKKQANDVSTLMCTPHYIHCINKPNEKRKRPD	597
OY		618	FDEVLIRIQVYITGLMENLRYVRRAAGFAVRKYEAFLQRYKSLCPETPMWAGRPDGVAV	677
Dd		598	WEESRVKQOIVEYLGIKENIRVRRAAGYARRIFQFKQLQRYALTITKATPSWOGEEKGYILH	657
OY		678	LVRHLGYPEEPYKMGRTKIFIIRPTLPATDISLEVBRQSALATKOAMRGFMHQFLR	737
Dd		658	LLOSVMNDSDPOFLGRSKVFTKAPESLFLBEMERKDIDYARIVLOKSMRFVARKKYYO	717
OY		738	VKRSALCIOSWMRGTLGRKKAARKMAQTIRLRIIGFILRHSPRCPENAEFLDHVASF	797
Dd		718	MR-----	719
OY		798	LNLRQDPRNVLDTSMTPPRALREASSELLRELCKMKMWKTCNISISEMKOLOOXAV	857
Dd		720	-----EEASDLL-----	726
OY		858	ASEIFGKKDNYPQSVPRLEFISTRIGTEISIPRYLOSLS-EPIDYAVPVYKDRKGXP	916
Dd		727	-----LKKKERRRSNSINNFEGIDVIOME-HPELOQFYGKREKIDFALTYVKYDRR-FKG	779
OY		917	RPROLLUTPSANVYVEDAKVQ-----RIDYANULTGISVSSLSDSLTFVLHVQ	963
Dd		780	VKROLLTLPKCLYLGIREKVQGGDPKGLVKEVLRKLEIEHILSYLSISTMODDIFILHEQ	839
OY		964	REDKKQGDVYLOGSDHYIELTKT---ALSADRYNNINIINGSTIFAG-----G	1009
Dd		840	-----EYDSLSEYKFKEEFLSLAKRREKTOQOLPKFSWTLELKLKENWG	887
OY		1010	PGRDGI-----IDFTSG-SELLITRAKNGHLAV-VAPRL--NSR	1044
Dd		888	PGVGAGSROYQHOGFGDLAVLKPMSKVLQVYSTIGPLPKNSR	930
<hr/>				
RESULT 11				
Q8WMN7 PRELIMINARY; PRT; 1098 AA.				
AC	O8WMN7;	01-MAR-2002	(Tremblrel. 20. Created)	
DT	01-MAR-2002	(Tremblrel. 20. Last sequence update)		
DT	01-JUN-2002	(Tremblrel. 21. last annotation update)		
DE	Myosin-1F.			
GN	MYOF.			
OS	Homo sapiens (Human).			
OC	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheta; Primates; Carnathini; Homnidae; Homo.			
OX	NCBI_TaxId=9606;			
[1]				
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21664126; PubMed=11804589;			
RA	Krugman S., Anderson K.E., Ridley S.H., Risco N., McGregor A.,			
RA	Coadwell J., Davidson K., Egutina G., Elison C.D., Ilip P.,			
RA	Manfava M., Klisckis N., Painter G., Thuring J.W., Cooper M.A.,			

QY 677 VLVRLHGYKPEEYKMGRTKIFIRPKTLFATEDSLEVRQSLATKIQAMRGFHNROKFL 736
 DB 655 HLLRAVMEPDQYOMGSTVYFVKNPESLLEVEERKKDGFARITQKAM----- 704
 QY 737 RKVRSACIQSWMRGTLGRKKAARMAAQTIRLLIRGFLIRHSPRCPENAFLDHVRAS 796
 DB 705 -----RRHVAVRKY----- 713
 QY 797 FLINLRQLPRNVLDTSWPTPPALREASBELRELCKMKNWYKCSISEPKKQOLQOKA 856
 DB 714 -----EEMRE----- 718
 QY 857 VASEIFKCKNDYPOSVPRLFTSTRIGTEEISPRVLQSLG-SEPLOYAVPVYKDKGK 915
 DB 719 EASNILNKKERRRRNSINRNFVDYGLBE-REPLKQFLGKKEKRVDFADSVTKYDKR-FK 776
 QY 916 PREPOLLTPSAVYVEDAK-----VKORIDYANLTGISVSLSDSLFVLHV 962
 DB 777 PIKRDILTPKCYVYIGREKMKGPBKQGYCEVLKKNVDIQALRGVSLSTRQDDFFIL-- 834
 QY 963 QREDNKOKGVYIQLSHVYETLTKT--ALSADRVANNINOGSTFFA----- 1007
 DB 835 -QED-----AADSFLSEYVKEFTEFVSLCKRFEENATRRPLTFSDTLQFRYKKEG 884
 QY 1008 GGPGRGDIIDFTSGSELLITKAKNGHLAV 1036
 DB 885 GGGGTSVTFRSRGFDLAVLKVGGRITLV 913

RESULT 12

063356 PRELIMINARY; PRT: 1107 AA.

AC 063356;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE MYOSIN I heavy chain.
 GN MYR3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRACUE-DAMLEY;
 RX MEDLINE=95247829; Pubmed=7730414;
 RA Stoeffler H.E., Ruppert C., Reinhard J., Bahler M.;
 RT "A novel mammalian myosin I from rat with an SH3 domain localizes to
 Con A-inducible, F-actin-rich structures at cell-cell contacts.";
 RL J. Cell Biol. 129:819-830(1995).
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL: X74815; CA52815.1; -;
 DR HSSP: P08799; LIND.
 DR InterPro: IPR000048; IQ_region.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ: 1.
 DR Pfam: PF00063; myosin_head: 1.
 DR Pfam: PF00018; SH3: 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR ProDom: PD000066; SH3: 1.
 DR ProDom: PD000355; myosin_head: 1.
 DR SMART: SM00242; MYSC: 1.
 DR SMART: SM00326; SH3: 1.
 DR PROSITE: PS50002; SH3: 1.
 DR SH3 domain.
 SO SEQUENCE 1107 AA; 126826 MW; B9D8FB0CE0471A8 CRC64;

Query Match 31.7%; Score 1712.5; DB 11; Length 1107;
 Best Local Similarity 37.2%; Pred. No. 3.3e-109;
 Matches 395; Conservative 175; Mismatches 295; Indels 197; Gaps 23;

QY 28 GVQDFVLEENFTSEAFIENTRRPRENLTITYIGPVLSVNPYRDLOIYSROHMEYRG 87
 DB 20 GVDDVVLNLSKIT-ESSIVENLKKRMMDYIFTYIGSVLISVNPVKOMPYEKEIKEMYG 78
 QY 88 VSEYVPPHLAVADTVYRALRTERDQAVNISGSGAGKTEATKRLLOFYAETCPAPER 147
 DB 79 AAQYENPHIYALADSMYRNMIDRENQCIVILISGSGAGKTVAAKYISVRSVSGGK 138
 QY 148 GGAVALDRLQSPVLEAGNAKTLRNDSSRFQYMOYQFPKCAPVGGHLSVLLKSR 207
 DB 139 VOHVNDILQSPVLEAGNAKTVANNSSRFQYFQSPGPGGKISNLFLEKSR 198
 QY 208 VYHONGERNFVFFQLEGEETLRIGLERNPQSYLYVKQCAKVSISNDKSPMKV 267
 DB 139 VYMRNPGERSFHFIFQLLEGASPEKQSLGI-TSDMYYYLISGSYKVDIDDKRDFQE 257
 QY 268 MKRALSVIDTEDEVEDLISVASYVLIHGNIFPADEDSNAQVTTENOAKTLTLLGVEG 327
 DB 258 TLHANVIGIFSEEQTLVQIVAGILHGINIF-KEVGNVAVSEEFAPAVYLGINQ 316
 QY 328 TLRLRALTHRKILIA-----KGEELSPULDEQAAVARADAKAVYSRPTTWLVKINSIA 383
 DB 317 DLKELKLSRQMDKSGKSESIVHTLVNEQACTRDLAKALHARVDFLVDSINKAM- 375
 QY 384 SKDASPSWRSSTVYGLDIYGEVFOHNSFEQFCINCNELQOLFTELTAKSEOEYE 443
 DB 376 EKDHEEYN-----IGVDIYGEFLFOANGFEQFCINVNNKIQIFELTKMAEQEYV 429
 QY 444 AEGIAMPEVOYFNNKIICDLVEKPK--GIISILDECL--RGEATDLFLKLEDTV 498
 DB 430 QEGIMWTFLEYNNKIICDLVEKPK--GIISILDECL--RGEATDLFLKLEDTV 489
 QY 499 KPHPHFLHKLADQTRKSLDGRFERLHYAGEVYVSTYGLDKNNDLFENLKETMSS 558
 DB 490 GSHEHF-----NSMNG-FTHHYGKSYDMDCGERRRDVLFMDLIEIMQSS 537
 QY 559 NNPIMAOCEFKSELSDKR-REPVTATQPKMSLQVLEILRKEPAYIRICIRPNADKOPGR 617
 DB 538 ELPIKLSLEPENLQADKKGRPTTAGSKIKKQANDLVSTLMCTPHYIRICIRPNKPKPD 597
 QY 618 FDEVLIRHOVYKLGIMENLRVARAGFAYRKRYEALQYKSKSCPETWMMAGRQDDGAV 677
 DB 598 WESEYKQHVEYLIGKENIRVARAGYAAVRVFOKFLQYALITLATWVMGDEKQVGLH 657
 QY 678 LVRLHGYKPEEYKMGRTKIFIRPKTLFATEDSLEVRQSLATKIQAMRGFHNROKFL 737
 DB 658 LIGSVNMSDQFOLGRSKYFIAPESTLLEEMREKKTGYARVYQKTWRKFVARKKYQ 717
 QY 738 VKRSAICIQSWMRGTLGRKKAARMAAQTIRLLIRGFLIRHSPRCPENAFPLDHVRAS 797
 DB 718 MRED----- 721
 QY 798 FLINLRQLPRNVLDTSWPTPPALREASBELRELCKMKNWYKCSISEPKKQOLQOKAV 857
 DB 722 ----- 721
 QY 858 VASEIFKCKNDYPOSVPRLFTSTRIGTEEISPRVLQSLG-SEPLOYAVPVYKDKGK 916
 DB 722 ASDLLNKKERRRRNSINRNFVDYGLBE-REPLKQFLGKKEKRVDFADSVTKYDKR-FKG 779
 QY 917 RPROLLTPSAVYVEDAKVQ-----RIDYANLTGISVSLSDSLFVLHVQ 963
 DB 780 VKRDLLTPKCYLILIGREKVGKPGDKVKEVLRKRIEVERLISVSLTMDIDIFILHEQ 839
 QY 964 REDNKOKGVYIQLSHVYETLTKT--ALSADRVANNINOGSTFFA-----G 1009
 DB 840 -----EYDLSLEVFKTEFLSLTRKYEETKQKPLAFSTLDELKAKKENWG 887
 QY 1010 P---GRDGIIDFTSG-SELLITKAKNGHLAV-VAPRL--NSR 1044
 DB 888 PMSAGSKROYQYQFGDLALIKPSNKVLYQVSIQGLPKNNR 929

RESULT 13

09NGUL7 PRELIMINARY; PRT; 1053 AA.
 AC 09NGUL7;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-DEC-2000 (Tremblrel. 15, Last sequence update)
 DE Amoeboid myosin I (Fragment)
 OS Strongylocentrotus purpuratus (Purple sea urchin)
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 NC NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20440049; PubMed=10984411;
 RA Strockin V., Seipel S., Krendel M., Bonder E.M.;
 RT Characterization of sea urchin unconventional myosins and analysis of
 RT their patterns of expression during early embryogenesis.
 RL Mol. Reprod. Dev. 57:111-126(2000).
 DR EMBL: AF248488; AAF71717.1;
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head.
 DR SMART: SM00242; MYSC; 1.
 FT NON_TER 1053 1053
 SQ SEQUENCE 1053 AA; 119723 MW; 0B31468B9F9FD03C CRC64;

Query Match 31.4%; Score 1695; DB 5; Length 1053;
 Best Local Similarity 37.4%; Pred. NO. 4.9e-108;
 Matches 400; Conservative 171; Mismatches 291; Indels 208; Gaps 24;

24 RDVGVQDFVLENTSEAFIENLRRENFLLTYTGVPVSNRYDQIYSRQME 83
 10 RKTSGEDMVLISKI-QGCAIYENIKKFMDDLTYTIGVPLISINPKMPFTEKEVD 68
 84 RRGVSEYVPPHLEFAVDYVRLKTERDQAVMISGESGAGTEATKRLQFYATCP 143
 69 MYGGAIIYENPPHYALTDYMYRMNLYEYNGCVIISGESGAGTAAKIMAYIAKV-- 126
 144 APERGA---VROLLQSNVPLFAFGAKTLRDNSSRFGKWDVDFEKGAPVGCHL 199
 127 ---SGGATVQVHKETIQSNPLEAFGNAKTVRNNSRFGKYVEIDFTKAGAPDGKIT 184
 200 SYLLEKSRVHQNGERNFHFYOLLEGEEETLRRLGRNPOSYLYLVKGCACAVSS 259
 185 NFLLEKSRVVSQNTDERNFHFYOLLEGEEETLRRLGRNPOSYLYLVKGCACAVSS 243
 260 NDKSDMKYMKALSYIDTEDEVLDLSIVASVLIHGINFPADEDSNAQVTTENOLKYL 319
 244 DDVEYKDTLNAHSVVISSEDDQENVLSVAGLIHGINVF-VEKDNVAVIHDEFLDPP 302
 320 TRLGVEGTLEALTRKIIA---KGEELSPLEAOAAVARDALAAVYRTTWLY 375
 303 SYLGLDKEALRKILSRNDKSKGAAEHIEVTLTEQAAHRRDLAAVHSRLDFLY 362
 376 RKIRSLASKDAESPMSRSTTVGLLDIYGFVEFHNSFEQFCINCEKLOLFTFLYL 435
 363 SSINGAKOKQVE-----ITIGVLDIYGFIFQNGFPOFCINVENKLOLFTFLYL 415
 436 KSEOEYEAGIAMEPYQENNTICDLYEERF-KGIISLDECL---RPGANDLFTL 491
 416 KAEDEEYVAEGIKKAKIDYFNNTICDLSKVPPEVMTCLDVCATMAVSGADEKLL 475
 492 EKLEDTVPRPHFLTKLADQKTRKSLDRGEFRLHYAGEVTVSVGFCDKNNDLFRLL 551
 476 QKMSAVGTHQHY-----GVLSGFLVHHYAAQVRYTVGFEKRRDVLFTDL 523
 552 KETWCSSMNPIMACF-DKSELSDKRPETVATQFMSLDQVLEILRSKAPAYIRICIPN 610

Db 524 IELMOSSSENFMTLPPEKILAGQAKSRPTTASSKIKTQANKLVKLMQCTPHYIRICIPN 583
 Qy 611 DAKOPGFDEVLNRHQVTKIGLMEINRVRRAGEFARRKYEAFLOKYSKLPETWPMAGR 670
 Db 584 ETRPQDEKERYHNOVEYLGLKENIRVRAEFARREDFKRLRAVAILTETWPMRGD 643
 Qy 671 PDGVAVLVRLHLYGKPEEYKMGRTKIFIRPTLTATEDSLVRROSLATKIQAMRGFH 730
 Db 644 VKSGVNHLLRAVMMDDERFOLGKTQVFTIKRPSLLEBQRERKDFHARIIOKAFQFN 703
 Qy 731 WRKFLRVKRSALCISQSMWRGTLGRRKAAKRRMAAQTIRRLTGFILHNSPCEMAFL 790
 Db 704 AQKYRLRLKQO-----
 Qy 791 DHVRASFLNLRQLPRNVLDTSWPTPPALREASELLRELCMKMVKYCRSISPEKQ 850
 Db 715 -----
 Qy 851 QLOQKAVASEIRKKNYPOSVPRLFTSTRLGT-EISPVYLSIGS--EPYQAVPVY 907
 Db 715 -----ASDILFDKERRRYSLNRFVGDYIGFEHNS--LQALVARRRIEFADKN 764
 Qy 908 KYDRKGRPRRQLLTPEAVYV--EDAK-----VKQRIDYANLTGISVSSLS 954
 Db 765 KYDRR-FKTKKDLISGRNITLIGREQIKKGPEKQIRDYVKKRIPLNHISVSLTLQ 823
 Qy 955 DSLFVLHVQREDNKQGVVLSQSDHVIETL---KTALSADRVNNINING----- 1002
 Db 824 DMFVLHVSEE-----YDLSLESTKTEFLVLVKNKRPETIQOLAMNENDSIFSVKKE 878
 Qy 1003 -----SIFPAGRGNDGIIDFTSGSELLITAKNGHLAVAPRLNSR 1044
 Db 878 GNGGCGTRVPRAGOGDPTAIL--KPSSKILTVSIGS-----LPR-NSR 920

RESULT 14
 061080
 ID 061080 PRELIMINARY; PRT; 1186 AA.
 AC 061080;
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Myosin IC heavy chain.
 GN MICHC.
 OS Acanthamoeba castellanii (Amoeba).
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.
 NC NCBI_TaxID=5753;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88016163; PubMed=3477803;
 RA Jung G., Korn E.D., Hammer J.A. III.
 RT "The heavy chain of Acanthamoeba myosin IB is a fusion of myosin-like
 RT and non-myosin-like sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 84:6720-6724(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99079990;
 RA Wang Z.Y., Wang F., Sellers J.R., Korn E.D., Hammer J.A. III.
 RT "Analysis of the regulatory phosphorylation site in Acanthamoeba
 RT myosin IC by using site-directed mutagenesis."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:15200-15205(1998).
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL: AF051353; AAC98089.1;
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head.
 DR SMART: SM00242; MYSC; 1.

SMART: SM00326; SH3: 1.
DR PROSITE; PS50002; SH3: 1.
KW SH3 domain.
SQ SEQUENCE 1186 AA; 129459 MW; E37AD44A685803A6 CRC64;

Query Match 29.18; Score 1573.5; DB 5; Length 1186;
Best Local Similarity 35.68; Pred. No. 1.4e-99;
Matches 373; Conservative 165; Mismatches 304; Indels 207; Gaps 23;

QY 28 GVQDVFLLNFTSEAFIENLRFRRENLIYTYIGPVLYVNPYRDL-QIYSQHMERFR 86
DB 8 GVDDVMVLISND-AINDMLKRRFADLIYTYIGHVLISVNPQIINNLTETETKDR 66
QY 87 GVSFEYEPHLEFAVDYVYRALFTRERDOAVMISGESGAGTEATKRLQFYAETCPAPE 146
DB 67 GKRYELDPHYALADDMYRMLSESDQCYIISGESGAGTEATKRLQFYAETCPAPE 126
QY 147 RGAAYRDLQSNPYLEAFNAKTLRNDNSSREKRYNDVDFGAVGHIISYLEKS 206
DB 127 DVMRKVYILSNPLLEAFNAKTLRNNSSREKRYNDVDFGAVGHIISYLEKS 186
QY 207 RVVHONHGERNFHYVYQLLEGEDETLRLGLEENPOSYLYVGCQAKYSINDKSDMK 266
DB 187 RVVYOTNGERNFHYVYQLLEGEDETLRLGLEENPOSYLYVGCQAKYSINDKSDMK 245
QY 267 VMKRALSYIDTEDEVEDLSIVASVHLGNHFADEDSNAQVYTTENQKLYTLLGVE 326
DB 246 DTVMAMKVIYGTAEQHEIFRLYVAYLYGNVQVDDCKGGSTIADKQVEMLAYLMKTE 305
QY 327 GTTLREALTHKRIAKGEE-----LSPILNLOQAVARALAKVSRFTVLRKIN 379
DB 306 PVAVQALILYR-TITTEGQGRSSVSCPDPLGATYKRALSLRNFQYTIQRYN 364
QY 365 DAMYIDDEPA-----LTGILDFEGEILFGNGEQLCNFVNEKLIQOIFQLTKAEQ 418
DB 440 EYEAEAGIAMEPVQYFNKKIICDLVEEAF-KGIISILDECLRPQADLFLTELEDTV 498
QY 419 EYEGAEIGIOMENIDFNKKICODLLEKRPGLMILITDVCFP-KGTDDEKRELELGA 477
DB 499 KHPHFLTKLADQKTRKSLDRGEFLHAYGEVYVYTGFLDKNNDLIRNLKETMCS 558
QY 478 PTHALHAATSQPD-----EVIHNGADVYVYNDGFDKKDLFLDLIGLACT 527
DB 559 MNPIMAGCFD--KSELSDKKRPETVATQKSLDLOVELLSKEPATYRCIKPNDAKPG 616
QY 528 STTFAGLFPPEAKKPTTAGFKIKESINILVATLSKTPHYTRCIRKPNKKAAN 587
DB 617 RDEVLIRHQVYIGLMENTLRYRAGFAFRKYEAFLOKYSILCEPTPMNAGRPODGA 676
QY 588 DPNNSLVHGVYTLGLEENVRIRAGYARQYNDKFEFRYRVYVCKTSMGNMGDVGAE 647
DB 677 VLVRLGYK-PEEYKMGRTIFIRPKTLEFATEDSLERQSLAKRIQAMGFMWRQK 735
QY 648 AILNHGMSLGEYQGGKTIKIPESVFSLEELDRITVFSYANKIRFLKRTAMRY 707
DB 736 LKVRSAIGIQSMWRGTGLRRKRAKRMKAQAOTIRLIRGLRHSRCPENAFILDHVA 795
QY 708 YEYKKG-----NDALVKKERRR----- 726
DB 796 SFLLNLRQULPRNVLDTSWTPPPALREASELLRELCKMNMWYKYSISPWKQOLQ 855
QY 727 ---LSLEP----- 732
DB 856 AVASEIFKGGKNTYPOSVRLPISLTGTEISPRVLOSSEPIQYAVPVYKDK--G 913
QY 733 ---KFDYIYRNFK--LKDCIGK-----GTEKVFALCKNNLDSEFG 773
DB 914 YKRPROLLTPSAVIVE-----DAKV-----KQIDYANLIGISVSLDLSL 958
QY 774 SKVERRIKVLISNMLVAIDPKKDIKKYKVPFLYVLRKRIDFNKIGSITLSPQDNM 833

QY 959 VLVHVEDNKKQGDVYVLOSDEVIETLTKAL-----SADRVNININGSI 1004
DB 834 LLSVNGEHS-----NLLECRKRELLGVLLKHNPSYRIQFADFN-----V 874
QY 1005 TRAGG-----PGRDGIIDFTSGS 1022
DB 875 TLKGGKTCVYKFRDPQGDGKVKGTYS 903

RESULT 15

ID Q19901 PRELIMINARY; PRT: 1100 AA.

AC Q19901; 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE F29D10.4 protein.

GN F29D10.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RA SEQUENCE FROM N.A.

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

RP MEDLINE=99069613; PubMed=9851916;

RA none;

RT Investigating biology.

RL Science 282:2012-2018(1998).

CC 1. SIMILARITY: CONTAINS 1 SH3 DOMAIN.

DR EMBL: Z75952; CAB00095.1; --

DR HSSP: P08799; IAMD.

DR InterPro: IPR000048; IQ_region.

DR InterPro: IPR001609; myosin_head.

DR InterPro: IPR001452; SH3.

DR InterPro: IPR000169; SH3prot_acsite.

DR Pfam: PF00612; IQ_1.

DR Pfam: PF00063; myosin_head_1.

DR Pfam: PF00018; SH3_1.

DR PRINTS: PR00193; MYOSINHEAVY.

DR ProDom: PD000066; SH3_1.

DR ProDom: PD000355; myosin_head_1.

DR SMART: SM00326; SH3: 1.

DR PROSITE: PS50002; SH3: 1.

DR PROSITE: PS00639; THIOLEPROTEASE_HIS; UNKNOWN_1.

DR SH3 domain.

SQ SEQUENCE 1100 AA; 124885 MW; 2CDB6F02EDEDED9 CRC64;

Query Match 29.08; Score 1564; DB 5; Length 1100;
Best Local Similarity 34.78; Pred. No. 5.7e-99;
Matches 367; Conservative 172; Mismatches 320; Indels 200; Gaps 21;

QY 27 GVQDVFLLNFTSEAFIENLRFRRENLIYTYIGPVLYVNPYRDL-QIYSQHMERFR 86
DB 14 GVDDVMVLPLRLT-EQSIENLKKRRLQANSITFYIGVLLSVNFKQMPYFTEKEMLLY 72
QY 87 GVSFEYEPHLEFAVDYVYRALFTRERDOAVMISGESGAGTEATKRLQFYAETCPAPE 146
DB 73 GAOYENAPHYALADDMYRMLSDNESQCYIISGESGAGTEATKRLQFYAETCPAPE 132
QY 147 RGAAYRDLQSNPYLEAFNAKTLRNDNSSREKRYNDVDFGAVGHIISYLEKS 206
DB 133 KYVHDKVILQSNPLLEAFNAKTLRNNSSREKRYNDVDFGAVGHIISYLEKS 192
QY 207 RVVHONHGERNFHYVYQLLEGEDETLRLGLEENPOSYLYVGCQAKYSINDKSDMK 266
DB 193 RVVHONHGERNFHYVYQLLEGEDETLRLGLEENPOSYLYVGCQAKYSINDKSDMK 251

```

QY 267 VMKALSVIDFTEDEVEDLISIVASVILHGNHFAADESDSMAQVTTENOLKYLTRLGVE 326
Db 252 STLHAKVGVNDQDQLEVLRIYAVTLHIGNTF-TEENNFAVSGDKYLEXPAPLGLT 310
QY 327 GTTLREALTHRIKIAK-----GELLSPNLNEQAAIYARDALAKAVSRTFTWLVRKINRSL 382
Db 311 SADIEAKLTGKKESKWKGTOKKEIDMKLWNEQASYTRDAMVKAIVARLEFDYLKKVNDAM 370
QY 383 ASKDAESPMSRSTTVLGLDIYGFVFNHNSFQFCINYNCEKLOOLFTELTKSQQEY 442
Db 371 ---NITSQSTSDNPSVGLDIYGFELFNNGFQFCINYNCEKLOOLFTELTKAQEY 427
QY 443 EABGIANEPYQYNNKTIICDLVEKF-KGIISILDECDL---PGEATDLTELEKLEDTV 498
Db 428 VREGIKWTEIDYEDNIVCDLIEKTRPPGIMSLDDTCAQNHQREGVDRQLTLTKSF 487
QY 499 KPHHFLTHKLAQDKTRKSLDRGEFRLHYAGEVTVSVYGFGLDKNDLFRNLKTCSS 558
Db 488 AGHPHF-----GPGSDSEYKHYAGDVYVNDGFCDRNRDVLYPDILIMOKS 535
QY 559 MNPIMACEDKS-ELSDKRREPETVANOQFMSLQLEILRSKEPAYIRICIRPNDAKOPR 617
Db 536 SRPFIOALFEPENVAASAGKRPJTSTFKIRTOANTLIVESLMKSPHYVRCIKNETKRPND 595
QY 618 FDEVILRHQVYKYLGLMENLRVRAAGAYRRTKEALQRYKSLCPETWPMAGRPQGVAV 677
Db 596 WEBSRVKHOVEYGLRENIYVRRAGAFYRPFADKFAORAYIVSPQWPCFQDQORACEI 655
QY 678 LVHNLGYPPEYVMGRTKIFIRFPKTLFATEDSLFVRQSLATKIOAMRGPHMRQKFLR 737
Db 656 ICDSVHMEKNQYOMGKITKIFVKNPESLFLLEETREKFDGYARVIOKAMROPSAR----- 710
QY 738 VKRSAICIOSMWRGTLGRKKAARKWAQTIIRLLRGFTLRHSPRCPENAFELDHVRASF 797
Db 711 ----- 710
QY 798 LNLRLRQLPRNVLDTSWTPPPALREASELLRELCKMNMWKCISPEWKQOLOQKAV 857
Db 711 -----KOHIKOEQ 719
QY 858 ASEIFGKKNDYPOSVPLFLISTRLGTEISPRVLSL--GSEPIOVAVPVVKYDRKGYK 915
Db 720 AADLMYKKERRRYSLNRFVGDYIGLEH--HPTLOSLVGRKQVLPACTANKYDKR-FR 776
QY 916 PRROLTLTPSAVIVVEDAKV-----KQIDYANLTGISVSLSDSLFVLHV 962
Db 777 VTKLDLLTVNHLTLIGKEKYNKNGPEKGIYEVIRQFDLPQIKSIGSPYQDDFVITL 836
QY 963 OREDNKQKGDVYLOS DHVIELTK---TALS---ADRYNN-----ININGST 1004
Db 837 GNDYSS-----LETPFKTEFCTALSKAYKERTNGTLHLDFRSSHVYKMKMF 886
QY 1005 TPAGSP-----GRDGIIDFTSGSELLITKANKHLLAVYA 1038
Db 887 DFSDGKRTYVGFENGD---TSSAEK--TLKPNKVLNVS 919

```

Search completed: July 7, 2003, 14:23:55
 Job time : 114.37 secs

•
•
•
•

	XX	11	48	61.5	1063	22	AAG3991	Human polypeptide
	XX	12	45	57.7	357	21	AAI6570	Arbidops thali
	XX	13	44	56.4	265	22	AAU36258	Pseudomonas aerug
	XX	14	44	56.4	268	18	AAW9839	L. pneumophila fa
	XX	15	41	52.6	415	23	ABR53255	Herbicidally acti
	XX	16	41	51.3	526	10	AAAP0587	Polypeptide with
	XX	17	40	51.3	33	22	AAO11660	Human polypeptide
	XX	18	40	51.3	46	22	AAO11393	Human polypeptide
	XX	19	40	51.3	51	23	ABP05815	Human polypeptide
	XX	20	40	51.3	402	23	AAU76528	Human OREX protei
	XX	21	40	51.3	402	23	AAO09883	Novel human secre
	XX	22	40	51.3	425	22	AAAS2493	Escherichia coli
	XX	23	39	50.0	64	22	AAO08242	Human polypeptide
	XX	24	39	50.0	70	22	AAU27464	Novel bone marrow
	XX	25	39	50.0	70	22	AAO08315	Zea mays protein
	XX	26	39	50.0	183	21	AAAS8516	Corn dihydrodipic
	XX	27	39	50.0	224	20	AAW97727	Neisseria gonorrh
	XX	28	39	50.0	261	21	AAAB03632	Neisseria gonorrh
	XX	29	39	50.0	261	21	AAAB03635	Neisseria gonorrh
	XX	30	39	50.0	261	21	AAAB03635	Neisseria gonorrh
	XX	31	39	50.0	261	21	AAAB03636	Neisseria gonorrh
	XX	32	39	50.0	262	23	ABR48921	Listeria monocyoc
	XX	33	39	50.0	279	21	AAG40867	Zea mays protein
	XX	34	39	50.0	315	21	AAAG40866	Zea mays protein
	XX	35	39	50.0	325	21	AAAG40865	Zea mays protein
	XX	36	38	48.7	2392	21	AAAB07665	Amino acid sequen
	XX	37	38	48.7	41	22	AAO13380	Human polypeptide
	XX	38	38	48.7	46	22	AAO09289	Human polypeptide
	XX	39	38	48.7	88	22	AAO09624	Human polypeptide
	XX	40	38	48.7	123	21	AAAG47869	Arabidopsis thalli
	XX	41	38	48.7	144	21	AAAG47868	Streptococcus pneu
	XX	42	38	48.7	157	19	AAAG62720	Arabidopsis thalli
	XX	43	38	48.7	170	21	AAAG47867	Arabidopsis thalli
	XX	44	38	48.7	218	21	AAAG14357	Arabidopsis thalli
	XX	45	38	48.7	219	21	AAAG47843	Arabidopsis thalli
	XX			48.7	239	21	AAAG14356	Arabidopsis thalli

ALIGNMENTS

RESULT 1	
ID	AAZ5521 standard; Protein; 262 AA.
AAZ5521	
AC	AAZ5521:
DT	30-SEP-1999 (first entry)
DE	E. coli mutant FabI protein.
KX	FabI; enoyl ACP reductase; acyl carrier protein; antimicrobial; therapy;
KW	fatty acid biosynthesis; detection; antibacterial; antifungal; treatment;
KM	antiprotozoal; antiparasitic; antiviral; soap; infection; disinfectant;
KV	detergent; mouthwash; toothpaste; contraceptive; inhibitor; tricosan.
XX	resistance; NSAM; non-specific antimicrobial; mutant.
OS	Escherichia coli.
XX	Synthetic.
Key	
FT	Location/Qualifiers
FT	1..262
FT	/label= FabI
FT	/note= "No stop codon given in specification"
PN	WO937800-A1.
PD	29-JUL-1999.
PE	22-JAN-1999; 99WO-US01288.
PR	26-JAN-1998; 98US-0013440.
RR	23-JAN-1998; 98US-0072244.

PA (TUFT) TUFTS COLLEGE.
 XX
 PI Levy SB, McMurry LM;
 XX
 DR WPI: 1999-458702/38.
 DR N-PSDB: AAX88348.
 XX
 PT Identifying antimicrobials that target enoyl acyl carrier protein
 PT reductase, for therapeutic use and for incorporation into e.g. soaps
 PT and detergents
 XX
 PS Claim 77; Fig 2; 80pp; English.
 XX
 CC This invention describes a novel method for identifying antimicrobials
 CC (1) that interact with an enoyl acyl carrier protein (ACP) reductase
 CC (ER) polypeptide, which is essential for fatty acid biosynthesis, by
 CC contacting ER with test compound and detecting any interaction. The
 CC identified antimicrobials have antibacterial, antifungal, antiprotzoal,
 CC antiparasitic and antiviral activity. The products of the invention are
 CC used (1) therapeutically to treat a wide variety of viral, bacterial,
 CC fungal, yeast and protozoal infections, in both humans and animals, and
 CC (11) in disinfectants, detergents, soaps, mouthwashes, toothpaste's and
 CC also contraceptive devices. The invention describes methods for (1)
 CC identifying a (1) from its effect on enzymatic activity of ER, (2)
 CC identifying a (1) from its ability to inhibit fatty acid biosynthesis in
 CC a microbe, (3) identifying a (1) that interacts with a mutant ER, (4)
 CC identifying (1) that inhibit proliferation or viability of a microbe that
 CC is resistant to triclosan or to NSAM (non-specific antimicrobials). This
 CC sequence represents a mutant Escherichia coli FabI protein which
 CC interacts with the enoyl-acyl carrier protein (ACP) reductase to
 CC illustrate the method of the invention.
 XX
 SQ Sequence 262 AA:
 XX
 Query Match 62.8%; Score 49; DB 20; Length 262;
 Best Local Similarity 60.0%; Pred. No. 2.2;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVRV 15
 :|||:|:|:|:|:|
 DB 170 VRYMANAMGPEGVRV 184
 XX
 RESULT 2
 AAU34559
 ID AAU34559 standard; Protein: 262 AA.
 XX
 AC AAU34559;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE E. coli cellular proliferation protein #140.
 XX
 KM Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 XX
 OS Escherichia coli.
 PN WO200170955-A2.
 PD 27-SEP-2001.
 PF 21-MAR-2001; 2001WO-US09180.
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.

XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR N-PSDB: AAS52418.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprising sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 10152; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 262 AA:
 XX
 Query Match 62.8%; Score 49; DB 22; Length 262;
 Best Local Similarity 60.0%; Pred. No. 2.2;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVRV 15
 :|||:|:|:|:|:|
 DB 170 VRYMANAMGPEGVRV 184
 XX
 RESULT 3
 AAE02203
 ID AAE02203 standard; Protein: 262 AA.
 XX
 AC AAE02203;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Escherichia coli NADPH-dependent enoyl-ACP reductase (FabI).
 XX
 KM FabI: high throughput method; fatty acid biosynthesis; therapy;
 KM bacterial enzyme; biological agent screening; otitis media; empyema;
 KM bacterial tracheitis; acute epiglottitis; thyroiditis; lung abscess;
 KM infective endocarditis; secretory diarrhoea; splenic; retroperitoneal;
 KM intrarenal; perinephric; cerebral; cutaneous; abscess; blepharitis;
 KM conjunctivitis; keratitis; endophthalmitis; preseptal; orbital;
 KM cellulitis; dacryocystitis; epididymitis; abscess; toxic shock syndrome;
 KM impetigo; folliculitis; wound infection; bacterial myositis;
 KM septic arthritis; osteomyelitis; enoyl-ACP reductase;
 KM acyl carrier protein.
 XX
 OS Escherichia coli.
 PN WO200130988-A1.
 PD 03-MAY-2001.
 PR 26-OCT-2000; 2000WO-US29451.
 PR 27-OCT-1999; 99US-0161775.
 XX
 PA

XX (SMIR) SMITHKLINE BEECHAM CORP.
 PA (SMIR) SMITHKLINE BEECHAM PLC.
 XX
 PI Dewolf W, Kallender H, Lonsdale JT;
 DR WPI; 2001-316332/33.
 DR N-PSDB; AAD06214.
 XX
 PT High throughput method for screening for biological agents against
 PT fatty acid biosynthesis comprises contacting a bacterial enzymatic
 PT pathway with enzymes e.g. malonyl-CoA:ACP transacylase -
 XX
 PS Claim 1; Page 15; 94pp; English.
 XX
 CC The present invention relates to a high throughput method for screening
 CC biological agents affecting fatty acid biosynthesis, comprises
 CC contacting a bacterial enzymatic pathway with enzymes. The method is
 CC used for screening biological agents affecting fatty acid biosynthesis.
 CC Agonists and antagonists of Fab (fatty acid biosynthesis) are used to
 CC inhibit, prevent or treat diseases such as infections of the upper
 CC respiratory tract (e.g. otitis media, bacterial tracheitis, acute
 CC epiglottitis, thyroiditis), lower respiratory (e.g. empyema, lung
 CC abscess), cardiac (e.g. infective endocarditis), gastrointestinal (e.g.
 CC secretory diarrhoea, splenic abscess, retroperitoneal abscess), central
 CC nervous system (e.g. cerebral abscess), eye (e.g. blepharitis),
 CC conjunctivitis, keratitis, endophthalmitis, preseptal and orbital
 CC cellulitis, dacryocystitis), kidney and urinary tract (e.g.
 CC epididymitis, intrarenal and perinephric abscess, toxic shock syndrome),
 CC skin (e.g. impetigo, folliculitis, cutaneous abscess, cellulitis, wound
 CC infection, bacterial myositis) and bone and joint (e.g. septic
 CC arthritis, osteomyelitis).
 CC The present sequence is Escherichia coli MADPH-dependent
 CC enoyl-ACP (acyl carrier protein) reductase (FabI). In fatty acid
 CC biosynthetic pathway, dehydration by FabZ leads to trans-2-enoyl-ACP
 CC which is in turn converted to acyl-ACP by FabI.
 XX
 SO Sequence 262 AA:
 Query Match 62.8%; Score 49; DB 22; Length 262;
 Best Local Similarity 60.0%; Pred. No. 2.2;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVRV 15
 :|||:|:|:|:|:|
 Db 170 VRYMANMGPEGVAV 184
 RESULT 4
 AAU38231
 ID AAU38231 standard; Protein; 269 AA.
 XX
 AC AAU38231;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Salmonella typhi cellular proliferation protein #122.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Salmonella typhi.
 XX
 PN W0200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS56090.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 13824; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 269 AA:
 Query Match 62.8%; Score 49; DB 22; Length 269;
 Best Local Similarity 60.0%; Pred. No. 2.3;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVRV 15
 :|||:|:|:|:|:|
 Db 177 VRYMANMGPEGVAV 191
 RESULT 5
 AAY25520
 ID AAY25520 standard; Protein; 429 AA.
 XX
 AC AAY25520;
 XX
 DT 30-SEP-1999 (first entry)
 XX
 DE E. coli FabI protein fragment.
 XX
 KW FabI; enoyl ACP reductase; acyl carrier protein; antimicrobial; therapy;
 KW fatty acid biosynthesis; detection; antibacterial; antifungal; treatment;
 KW antiprotozoal; antiparasitic; antiviral; soap; infection; disinfectant;
 KW detergent; mouthwash; toothpaste; contraceptive; inhibitor; triclosan;
 KW resistance; NSAM; non-specific antimicrobial.
 XX
 OS Escherichia coli.
 XX
 PF Key Location/Qualifiers
 FT Protein 1..429
 FT /label= FabI
 FT /note= "No start or stop codon given in specification"
 FT MISC-difference 74
 FT /note= "in-frame stop codon encoded by TGA"
 FT MISC-difference 111

FT /note= "in-frame stop codon encoded by TAA"
 FT Misc-difference 393
 FT /note= "in-frame stop codon encoded by TAA"
 FT Misc-difference 398
 FT /note= "in-frame stop codon encoded by TAA"
 XX
 PN WO9937800-A1.
 XX
 PD 29-JUL-1999.
 XX
 PF 22-JAN-1999; 99WO-US01288.
 XX
 PR 26-JAN-1998; 98US-0013440.
 PR 23-JAN-1998; 98US-0072244.
 XX
 PA (TUFTS) TUFTS COLLEGE.
 PI Levy SB, McMurry LM.
 PI WPI: 1999-458702/38.
 DR N-PSDB; AAX88347.
 XX
 PT Identifying antimicrobials that target enoyl acyl carrier protein
 PT reductase, for therapeutic use and for incorporation into e.g. soaps
 PT and detergents
 PS Example 5; Page 68-70; 80pp; English.
 XX
 CC This invention describes a novel method for identifying antimicrobials
 CC (I) that interact with an enoyl acyl carrier protein (ACP) reductase
 CC (ER) polypeptide, which is essential for fatty acid biosynthesis, by
 CC contacting ER with test compound and detecting any interaction. The
 CC identified antimicrobials have antibacterial, antifungal, antiprotazoal,
 CC antiparasitic and antiviral activity. The products of the invention are
 CC used (i) therapeutically to treat a wide variety of viral, bacterial,
 CC fungal, yeast and protozoal infections, in both humans and animals, and
 CC (ii) in disinfectants, detergents, soaps, mouthwashes, toothpaste's and
 CC also contraceptive devices. The invention describes methods for (1)
 CC identifying a (I) from its effect on enzymatic activity of ER, (2)
 CC identifying a (I) from its ability to inhibit fatty acid biosynthesis in
 CC a microbe, (3) identifying a (I) that interacts with a mutant ER, (4)
 CC identifying a (I) that inhibit proliferation or viability of a microbe that
 CC is resistant to triclosan or to NSAM (non-specific antimicrobials). This
 CC sequence represents the Escherichia coli FabI protein which interacts
 CC with the enoyl-acyl carrier protein (ACP) reductase to illustrate the
 CC method of the invention.
 CC
 SQ Sequence 429 AA;
 XX
 XX Query Match 62.8%; Score 49; DB 20; Length 429;
 XX Best Local Similarity 60.0%; Pred. No. 3.7;
 XX Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MRRASALGSDGVRV 15
 :|||:|:|:|:|:|
 DB 300 VRYMANAMGPEGVRV 314
 RESULT 6
 ABG28562
 ID ABG28562 standard; Protein; 491 AA.
 AC ABG28562;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #28553.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX

PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB; AAS92749.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 20; SEQ ID NO 58921; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 491 AA;
 XX
 XX Query Match 62.8%; Score 49; DB 22; Length 491;
 XX Best Local Similarity 60.0%; Pred. No. 4.3;
 XX Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MRRASALGSDGVRV 15
 :|||:|:|:|:|:|
 DB 378 VRYMANAMGPEGVRV 392
 RESULT 7
 ABG28564
 ID ABG28564 standard; Protein; 693 AA.
 AC ABG28564;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #28555.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX

PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR N-PSDB: AAS92751.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 20: SEQ ID No 58923; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantifying a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcr_sequences.
XX
XX SQ Sequence 693 AA;
XX
XX Query Match 62.8%; Score 49; DB 22; Length 693;
XX Best Local Similarity 60.0%; Pred. No. 6.2;
XX Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 MRYRASALGSDGVRV 15
XX :|||:|:|:|:|
XX 601 VRYMANMNGPEGVRV 615
XX
XX RESULT 8
XX AAB90827
XX ID AAB90827 standard; Protein; 267 AA.
XX
XX AC AAB90827;
XX
XX DT 15-JUN-2001 (first entry)
XX
XX DE Human shear stress-response protein SEQ ID NO: 171.
XX
XX KW Human; shear stress-response protein; vascular disease;
XX arteriosclerosis.
XX
XX OS Homo sapiens.
XX
XX PN WO200125427-A1.
XX
XX PD 12-APR-2001.
XX

PF 02-OCT-2000; 2000WO-JP06840.
XX
XX 01-OCT-1999; 99JP-0280976.
XX
XX (KYOW) KYOWA HAKKO KOSYO KK.
XX (NOJ) NOJIMA H.
XX
XX Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
PI Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
XX WPI: 2001-266308/27.
DR N-PSDB: AAH02967.
XX
XX DNA sequences, proteins encoded by them and antibodies against them
PT useful in diagnosis and treatment of vascular disease caused by
PT arteriosclerosis -
XX
XX Claim 35; Page 664-665; 678pp; Japanese.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension.
XX
XX SQ Sequence 267 AA;
XX
XX Query Match 61.5%; Score 48; DB 22; Length 267;
XX Best Local Similarity 100.0%; Pred. No. 3.3;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 7 ALGSDGVRV 16
XX :|||||
XX 26 ALGSDGVRV 35
XX
XX DB
XX
XX RESULT 9
XX AAM41777
XX ID AAM41777 standard; Protein; 537 AA.
XX
XX AC AAM41777;
XX
XX DT 22-OCT-2001 (first entry)
XX
XX DE Human polypeptide SEQ ID NO 6708.
XX
XX KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX
XX OS Homo sapiens.
XX
XX PN WO200153312-A1.
XX
XX PD 26-JUL-2001.
XX
XX PF 26-DEC-2000; 2000WO-US34263.
XX
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB; AAI60933.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Example 2: SEQ ID NO 6708; 10076bp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SO Sequence 537 AA;
 Query Match 61.5%; Score 48; DB 22; Length 537;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 ALGSDGVRVT 16
 DB 60 ALGSDGVRVT 69

RESULT 10
 ID ABG10171 standard; Protein; 1050 AA.
 XX
 AC ABG10171;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #10162.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT.
 XX
 DR WPI, 2001-639362/73.
 DR N-PSDB; AAS74358.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 40530; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SO Sequence 1050 AA;
 Query Match 61.5%; Score 48; DB 22; Length 1050;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 ALGSDGVRVT 16
 DB 12 ALGSDGVRVT 21

RESULT 11
 ID AAM39991 standard; Protein; 1063 AA.
 XX
 AC AAM39991;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 3136.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0635450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QH, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB: AA159147.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PS Example 4; SEQ ID NO 3136; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA158642-AA162213) with neotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 1063 AA;
SO
Query Match 61.5%; Score 48; DB 22; Length 1063;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 ALGSDGVRVT 16
Db 26 ALGSDGVRVT 35
RESULT 12
ID AAG16520
XX AAG16520 standard; Protein; 357 AA.
AC AAG16520;
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 17197.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137522.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

```

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145227.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

```

```

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 57.7%; Score 45; DB 21; Length 357;
Best Local Similarity 69.2%; Pred. No. 15;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 4 RASALGSDGYRT 16
DB 43 RASALGNTGLKVT 55

```

```

RESULT 13
AAU36258
ID AAU36258 standard; Protein: 265 AA.
XX
AC AAU36258;
XX
DT 14-FEB-2002 (first entry)
XX
DE Pseudomonas aeruginosa cellular proliferation protein #248.
XX
KM Antisense; prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.
XX
OS Pseudomonas aeruginosa.
XX
PN W0200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELITR) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
XX N-PSDB: AAS54117.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -

```

PS Example 3: Seq ID No 11851; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 265 AA;

Query Match 56.4%; Score 44; DB 22; Length 265;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRY 15
DB 173 VRYLAGSLGAGETRV 187

RESULT 14
AAW99839
ID AAW99839 standard; Protein: 268 AA.

XX AAW99839;

AC AAW99839;

DT 08-JUN-1999 (first entry)

XX L. pneumophila Fabi enoyl reductase homologue protein sequence.

DE Legioneella pneumophila; Fabi; enoyl reductase; antimicrobial;
KW genfibrozil; growth inhibition; bacterium; infection.

XX Legioneella pneumophila.

OS

XX WO9731530-A1.

PN

XX 04-SEP-1997.

PD

XX 28-FEB-1997; 97WO-US03158.

PF

XX 29-FEB-1996; 96US-0608712.

PR

XX (UYCO) UNIV COLUMBIA NEW YORK.

PA

XX Della-Latta P, Kabbash C, Shuman HA, Silverstein SC;
PI WPI: 1997-448377/41.
XX N-PSDB; AAX19775.
DR

XX Inhibiting growth of bacteria - by contacting the bacteria with
PT genfibrozil or related compound

XX

PS Example 2; Fig 21B; 109pp; English.

XX A method has been developed for: (i) inhibiting growth of a bacterium;
CC (ii) alleviating the symptoms of a bacterial infection in a subject;
CC (iii) inhibiting activity of enoyl reductase enzyme in a cell; or (iv)
CC altering a pathway of fatty acid synthesis in a bacterium, comprises
CC contacting the bacterium, subject or cell with a novel compound (I) or

CC its salt or ester. The above processes are especially applicable to
CC bacteria/bacterial cells selected from *Legioneella pneumophila*,
CC *Mycobacterium tuberculosis*, *Bacillus subtilis*, *Bacillus megaterium*,
CC *Pseudomonas oleovorans*, *Alcaligenes eutrophus*, *Rhodococcus sp.*,
CC *Citrobacter freundii*, Group A *Streptococcus sp.*, *Coag neg Staphylococcus*
CC *aureus* or *Noctuidia sp.* The bacterial infection is e.g. leprosy, brucella
CC or salmonella. The present sequence represents a Fabi enoyl reductase
CC homologue from *L. pneumophila*, which is used in an example from the
CC present invention.

XX Sequence 268 AA;

Query Match 56.4%; Score 44; DB 18; Length 268;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRY 15
DB 177 VRYLAGSLGSRGLRI 191

RESULT 15
ABB93255
ID ABB93255 standard; Protein: 415 AA.

XX ABB93255;

AC ABB93255;

DT 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 2466.

DE Herbicidally active polypeptide SEQ ID NO 2466.

XX

XX Herbicidal; plant; agriculture; herbicide.

XX

XX Arabidopsis thaliana.

OS

XX WO200210210-A2.

PN

XX 07-FEB-2002.

PD

XX 28-AUG-2001; 2001WO-EP09892.

PF

XX 28-AUG-2001; 2001WO-EP09892.

PR

XX (FARB) BAYER AG.

PA

XX Tietjen K, Weidler M;

PI WPI: 2002-269010/31.

DR

XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -

XX

XX Claim 5; SEQ ID NO 2466; 261pp + Sequence Listing; English.

PS

XX The invention relates to identifying target proteins
CC aligning and comparing nucleic acid or amino acid sequences, comprising
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.

XX

XX Sequence 415 AA;

Query Match 52.6%; Score 41; DB 23; Length 415;
Best Local Similarity 53.3%; Pred. No. 82;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRY 16

Mon Jul 7 15:30:57 2003

us-09-893-371-2.rag

Page 10

DB 392 RYKAGALGAERKRAT 406

Search completed: July 7, 2003, 14:21:22
Job time : 3.53962 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:19:35 ; Search time 0.543396 Seconds
(without alignments)
866.342 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78
Sequence: 1 MRYRASALGSDGVRV 16

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.8	262	1	US-08-241-766-8
2	49	62.8	262	1	US-08-241-766-9
3	38	48.7	308	4	US-09-347-803-12
4	38	48.7	308	5	PCT-US96-03916-6
5	38	48.7	985	5	PCT-US96-03916-6
6	37	47.4	309	4	US-09-134-001C-3039
7	37	47.4	623	3	US-09-041-991A-6
8	37	47.4	943	4	US-09-397-885-5
9	36	46.2	260	4	US-09-134-001C-4009
10	36	46.2	537	1	US-08-173-508-2
11	36	46.2	537	2	US-08-265-310-2
12	36	46.2	537	3	US-08-951-742-2
13	36	46.2	635	3	US-09-041-991A-4
14	36	46.2	702	6	5240838-15
15	36	46.2	3729	2	US-08-804-227C-4
16	35	44.9	48	2	US-08-637-759B-480
17	35	44.9	48	3	US-08-871-355A-480
18	35	44.9	256	4	US-09-292-412-2
19	35	44.9	303	4	US-09-420-786A-3
20	35	44.9	308	4	US-09-347-803-10
21	35	44.9	338	4	US-09-199-637A-47
22	35	44.9	350	4	US-08-961-083-126
23	35	44.9	5588	4	US-09-036-987A-6
24	35	44.9	5588	4	US-09-370-700-6
25	34	43.6	77	4	US-09-465-558-52
26	34	43.6	113	1	US-08-211-202-114
27	34	43.6	113	1	US-08-211-202-114

28	34	43.6	303	2	US-08-286-819A-23	Sequence 23, Appl
29	34	43.6	303	3	US-08-980-357-23	Sequence 23, Appl
30	34	43.6	319	2	US-08-795-927-4	Sequence 4, Appl
31	34	43.6	329	3	US-09-022-669-2	Sequence 2, Appl
32	34	43.6	329	4	US-09-262-748-2	Sequence 2, Appl
33	34	43.6	329	4	US-09-603-567-2	Sequence 2, Appl
34	34	43.6	359	4	US-09-722-139-4	Sequence 4, Appl
35	34	43.6	359	4	US-09-721-832-4	Sequence 4, Appl
36	34	43.6	359	4	US-09-721-832-4	Sequence 4, Appl
37	34	43.6	433	4	US-09-199-637A-349	Sequence 349, App
38	34	43.6	524	1	US-08-529-654-4	Sequence 4, Appl
39	34	43.6	524	3	US-08-957-302A-2	Sequence 2, Appl
40	34	43.6	524	4	US-09-542-403-2	Sequence 2, Appl
41	34	43.6	524	4	US-09-542-403-2	Sequence 2, Appl
42	34	43.6	623	1	US-08-653-740-7	Sequence 7, Appl
43	34	43.6	623	2	US-09-073-594-7	Sequence 7, Appl
44	34	43.6	623	3	US-09-275-925-7	Sequence 7, Appl
45	34	43.6	633	3	US-09-041-991A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-241-766-8
; Sequence 8, Application US/08241766
; Patent No. 5686350
; GENERAL INFORMATION:
; APPLICANT: JACOBS, W. R.
; APPLICANT: COLLINS, D. M.
; APPLICANT: BANERJEE, A. M.
; APPLICANT: DELISLE, G. W.
; APPLICANT: WILSON, T. M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN LHA AGENT
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & ROESTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,766
; FILING DATE: 12-MAY-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, GLADYS H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 25237-20003.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-241-766-8
Query Match 62.8%; Score 49; DB 1; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.51;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
1 MRYRASALGSDGVRV 15

Db :|||:|:|:|:|
170 VRYMANMGPEGVRY 184

RESULT 2

US-08-241-766-9
Sequence 9, Application US/08241766
Patent No. 5686590
GENERAL INFORMATION:
APPLICANT: JACOBS, W. R.
APPLICANT: COLLINS, D. M.
APPLICANT: BANERJEE, A.
APPLICANT: DELISLE, G. W.
APPLICANT: WILSON, T. M.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN INHA AGENT
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/241,766
FILING DATE: 12-MAY-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MONROY, GLADIS H.
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 25237-20003.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-241-766-9

Query Match 62.8%; Score 49; DB 1; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.51;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MYRASALGSDGVRY 15
Db 170 VRYMANMGPEGVRY 184

RESULT 3
US-09-347-803-12
Sequence 12, Application US/09347803
Patent No. 6274379
GENERAL INFORMATION:
APPLICANT: Famodu, Layo O.
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Orozco, Buddy
TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes
FILE REFERENCE: BB-1176
CURRENT APPLICATION NUMBER: US/09/347,803
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: 60/092,952
EARLIER FILING DATE: July 15, 1998

NUMBER OF SEQ ID NOS: 27
SOFTWARE: Microsoft Office 97
SEQ ID NO 12
LENGTH: 308
TYPE: PRT
ORGANISM: Oryza sativa
US-09-347-803-12

Query Match 48.7%; Score 38; DB 4; Length 308;
Best Local Similarity 88.9%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ASALGSDGV 13
Db 120 ASALGDDGV 128

RESULT 4

PCT-US96-03916-6
Sequence 6, Application PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-03916-6

Query Match 48.7%; Score 38; DB 5; Length 985;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 YRASALGSDGVRY 16
Db 921 YECTVLISDGTRVT 934

RESULT 5
PCT-US96-03916-6
Sequence 66, Application PC/TUS9603916
GENERAL INFORMATION:

```

OY      2 RYRALSQSDGVRV 15
      11 :|||:|:|:
Db      272 RYRHTALGTDIPVKL 285

RESULT 7
US-09-041-991A-6
: Sequence 6, Application US/09041991A
: Patent No. 6107278
: GENERAL INFORMATION:
: APPLICANT: Schnepf, H. Ernest
: APPLICANT: Narva, Kenneth E.
: APPLICANT: Muller-Cohn, Judy
: TITLE OF INVENTION: Toxins Active Against Pests
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Salliwanchik, Lloyd & Salliwanchik
: STREET: 2421 N.W. 41st Street, Suite A-1
: CITY: Gainesville
: STATE: Florida
: COUNTRY: USA
: ZIP: 32606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/041,991A
: FILING DATE: 13-MAR-1998
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Sanders, Jay M.
: REGISTRATION NUMBER: 39,355
: REFERENCE/DOCKET NUMBER: MA-709
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (352) 375-8100
: TELEFAX: (352) 372-5800
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 623 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-09-041-991A-6

Query Match      47.4%; Score 37; DB 3; Length 623;
Best Local Similarity 53.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      4 RASALGSDGVRV 16
      11 :||:|:|
Db      539 RVSSIGSSTRV 551

RESULT 8
US-09-397-885-5
: Sequence 5, Application US/09397885
: Patent No. 6133007
: GENERAL INFORMATION:
: APPLICANT: Moller, Soren
: APPLICANT: Johansen, Charlotte
: APPLICANT: Schafer, Thomas
: APPLICANT: Ostergaard, Peter Rahbek
: APPLICANT: Heock, Lisbeth Hedegaard
: TITLE OF INVENTION: A 2,6,-b-D-Fructan Hydrolase Enzyme And
: FILE REFERENCE: 5540 200-US
: CURRENT APPLICATION NUMBER: US/09/397,885
: CURRENT FILING DATE: 1999-09-17
: EARLIER APPLICATION NUMBER: PA 1998 01173

```

EARLIER FILING DATE: 1998-09-18
EARLIER APPLICATION NUMBER: PA 1998 01623
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: 60/101,615
EARLIER FILING DATE: 1998-09-24
EARLIER APPLICATION NUMBER: 60/111,675
EARLIER FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 943
TYPE: PRF
ORGANISM: Paenibacillus macerans
US-09-397-885-5

Query Match 47.4%; Score 37; DB 4; Length 943;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 MYRASALGSDGV 12
DB 270 LRRRANGQGTGD 281

RESULT 9
US-09-134-001C-4009
Sequence 4009, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GRC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4009
LENGTH: 260
TYPE: PRF
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4009

Query Match 46.2%; Score 36; DB 4; Length 260;
Best Local Similarity 46.7%; Pred. No. 88;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MYRASALGSDGV 15
DB 175 VKYIALDLGDNIRV 189

RESULT 10
US-08-173-508-2
Sequence 2, Application US/08173508
Patent No. 5616485
GENERAL INFORMATION:
APPLICANT: Bartfield, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 508/173,508
APPLICATION NUMBER: US/08/173,508
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/125/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-173-508-2

Query Match 46.2%; Score 36; DB 1; Length 537;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MYRASALGSDGV 16
DB 6 IRRRAFATGALV 21

RESULT 11
US-08-265-310-2
Sequence 2, Application US/08265310
Patent No. 5856166
GENERAL INFORMATION:
APPLICANT: Bartfield, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
APPLICANT: Malek, Lawrence T.
APPLICANT: Soostmeyer, Gisela
APPLICANT: Walczyk, Eva
APPLICANT: Krygsman, Phyllis
APPLICANT: Garven, Sheila
TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 508/265,310
APPLICATION NUMBER: US/08/265,310
FILING DATE: 24-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,508
FILING DATE: 23-DEC-1993

APPLICATION NUMBER: 300,211
 FILING DATE: 23-JAN-1989
 APPLICATION NUMBER: 759,315
 FILING DATE: 26-JUL-1985
 SEQ ID NO:15
 LENGTH: 702
 5240838-15

Query Match 46.2%; Score 36; DB 6; Length 702;
 Best Local Similarity 77.8%; Pred. No. 2.6e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 SALSGDVR 14
 :||| |||
 Db 277 NALGEDVR 285

RESULT 15
 US-08-804-227C-4
 ; Sequence 4, Application US/08804227C
 ; Patent No. 5876991
 ; GENERAL INFORMATION:
 ; APPLICANT: Dehoff, Bradley S.
 ; APPLICANT: Rostock, Paul R., Jr.
 ; APPLICANT: Sutton, Kimberly L.
 ; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: THOMAS G. PLANT 1501
 ; STREET: LILLY CORPORATE CENTER
 ; CITY: INDIANAPOLIS
 ; STATE: IN
 ; COUNTRY: USA
 ; ZIP: 46285
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: ASCII(DOS) Text only
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/804,227C
 ; FILING DATE: February 21, 1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plant, Thomas, G.
 ; REGISTRATION NUMBER: 35,784
 ; REFERENCE/DOCKET NUMBER: X-8231
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 317-276-2459
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3729 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-804-227C-4

Query Match 46.2%; Score 36; DB 2; Length 3729;
 Best Local Similarity 66.7%; Pred. No. 1.6e+03;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RYRSALGSDGV 13
 :||| |||
 Db 1778 RSRAPAGADGV 1789

Search completed: July 7, 2003, 14:25:41
 Job time : 2.5434 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:21:30 ; Search time 0.664151 Seconds

(without alignments)
2770.284 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78

Sequence: 1 MRYRASALGSDGVRYT 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PC1_NEM_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PC105_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.8	262	10 US-09-815-242-10152	Sequence 10152, A
2	49	62.8	269	10 US-09-815-242-13824	Sequence 13824, A
3	44	56.4	265	10 US-09-815-242-11851	Sequence 11851, A
4	41	52.6	451	9 US-10-128-714-3277	Sequence 3277, Ap
5	41	52.6	512	9 US-10-156-761-8793	Sequence 8793, Ap
6	41	52.6	559	9 US-10-128-714-8277	Sequence 8277, Ap
7	39	50.0	1000	9 US-10-156-761-13499	Sequence 13499, A
8	38	48.7	308	9 US-09-882-691-12	Sequence 12, Appl
9	38	48.7	315	10 US-09-815-242-13289	Sequence 13289, A
10	38	48.7	985	9 US-09-994-064-6	Sequence 6, Appl
11	38	48.7	985	9 US-09-994-064-6	Sequence 6, Appl
12	38	48.7	1233	9 US-09-738-626-4312	Sequence 4312, Ap
13	37.5	48.1	355	9 US-10-156-761-15060	Sequence 15060, A
14	37	47.4	172	9 US-10-156-761-14091	Sequence 14091, A
15	37	47.4	206	9 US-10-156-761-10243	Sequence 10243, A
16	37	47.4	252	9 US-09-880-748-1583	Sequence 1583, Ap
17	37	47.4	320	9 US-10-156-761-13166	Sequence 13166, A
18	37	47.4	338	9 US-10-156-761-13572	Sequence 13572, A
19	37	47.4	441	9 US-09-738-626-4275	Sequence 4275, Ap

20	37	47.4	511	9 US-10-156-761-8583	Sequence 8583, Ap
21	37	47.4	530	9 US-10-156-761-10819	Sequence 10819, A
22	37	47.4	587	9 US-10-156-761-8493	Sequence 8493, Ap
23	37	47.4	722	9 US-09-738-626-5453	Sequence 5453, Ap
24	37	47.4	747	9 US-10-156-761-8661	Sequence 8661, Ap
25	37	47.4	943	10 US-09-969-362-5	Sequence 5, Appl
26	37	47.4	1225	9 US-10-156-761-12577	Sequence 12577, A
27	37	47.4	1316	9 US-10-120-544A-4	Sequence 4, Appl
28	37	47.4	1344	9 US-10-120-544A-4	Sequence 20, Appl
29	37	47.4	1386	9 US-10-120-544A-20	Sequence 40694, A
30	36	46.2	30	10 US-10-120-544A-6	Sequence 7, Appl
31	36	46.2	135	9 US-09-864-761-40694	Sequence 4818, Ap
32	36	46.2	187	9 US-10-137-077-7	Sequence 11279, A
33	36	46.2	295	10 US-09-815-242-11279	Sequence 8219, A
34	36	46.2	305	9 US-10-156-761-88219	Sequence 4829, Ap
35	36	46.2	358	9 US-09-738-626-4829	Sequence 9458, Ap
36	36	46.2	377	9 US-10-156-761-9458	Sequence 11514, A
37	36	46.2	438	9 US-10-156-761-11514	Sequence 234, Ap
38	36	46.2	1160	9 US-10-028-072-234	Sequence 234, Ap
39	36	46.2	1160	9 US-10-121-049-234	Sequence 234, Ap
40	36	46.2	1160	9 US-10-123-904-234	Sequence 234, Ap
41	36	46.2	1160	9 US-10-140-470-234	Sequence 234, Ap
42	36	46.2	1160	9 US-10-175-746-234	Sequence 234, Ap
43	36	46.2	1160	9 US-10-176-918-234	Sequence 234, Ap
44	36	46.2	1160	9 US-10-176-921-234	Sequence 234, Ap
45	36	46.2	1160	9 US-10-137-865-234	Sequence 234, Ap

ALIGNMENTS

RESULT 1
US-09-815-242-10152
Sequence 10152, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: ELITRA.011A
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 10152
LENGTH: 262
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10152
Query Match 62.8% Score 49; DB 10; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.59;

Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||||:|||||

Db 170 VRYMANMGPGEVR 184

RESULT 2

US-09-815-242-13824

Sequence 13824, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA 011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21, 078

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13824

LENGTH: 269

TYPE: PRT

ORGANISM: Salmonella typhi

US-09-815-242-13824

Query Match 62.8%; Score 49; DB 10; Length 269;

Best Local Similarity 60.0%; Pred. No. 0.61;

Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||||:|||||

Db 177 VRYMANMGPGEVR 191

RESULT 3

US-09-815-242-11851

Sequence 11851, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA 011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11851

LENGTH: 265

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-815-242-11851

Query Match

Best Local Similarity 56.4%; Score 44; DB 10; Length 265;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||||:|||||

Db 173 VRYMANMGPGEVR 187

RESULT 4

US-10-128-714-3277

Sequence 3277, Application US/10128714

Publication No. US20030119013A1

GENERAL INFORMATION:

APPLICANT: Jiang, Bo

APPLICANT: Hu, Wenqi

APPLICANT: Tishkoff, Daniel

APPLICANT: Zamudio, Carlos

APPLICANT: Eroshkin, Alexey M

APPLICANT: Lemieux, Sebastien M

TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

FILE REFERENCE: 10182-018-999

CURRENT APPLICATION NUMBER: US/10/128,714

CURRENT FILING DATE: 2002-04-23

PRIOR APPLICATION NUMBER: US 60/285,697

PRIOR FILING DATE: 2001-04-23

PRIOR APPLICATION NUMBER: US 60/287,066

PRIOR FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: US 60/295,890

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: US 60/303,899

PRIOR FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: US 60/316,362

PRIOR FILING DATE: 2001-08-31

NUMBER OF SEQ ID NOS: 8603

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3277

LENGTH: 451

TYPE: PRT

ORGANISM: Aspergillus fumigatus

US-10-128-714-3277

Query Match 52.6%; Score 41; DB 9; Length 451;

Best Local Similarity 60.0%; Pred. No. 31;

Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||||:|||||

Db 14 MROEAKATGKDGPRV 28

RESULT 5

US-10-156-761-9793
; Sequence 9793, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9793
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9793

Query Match 52.6%; Score 41; DB 9; Length 512;
Best Local Similarity 61.5%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YRASALGSDGVR 15
DB 136 YRVALDSKGRV 148

RESULT 6

US-10-128-714-8277
; Sequence 8277, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8277
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8277

Query Match 52.6%; Score 41; DB 9; Length 559;

Best Local Similarity 60.0%; Pred. No. 40;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
DB 138 MRQEAATGKGDGPRV 152

RESULT 7

US-10-156-761-13499
; Sequence 13499, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13499
; LENGTH: 1000
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13499

Query Match 50.0%; Score 39; DB 9; Length 1000;
Best Local Similarity 61.5%; Pred. No. 1,8e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RYRASALGSDGVR 14
DB 839 RYPAEGIGDDAVR 851

RESULT 8

US-09-882-691-12
; Sequence 12, Application US/09882691
; Publication No. US20030104593A1
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes
; FILE REFERENCE: BB-1176
; CURRENT APPLICATION NUMBER: US/09/882,691
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/092,952
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-882-691-12

Query Match 48.7%; Score 38; DB 9; Length 308;
Best Local Similarity 88.9%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ASALGSDGV 13
||||| |||

Db 120 ASALGDDGV 128

RESULT 9

US-09-815-242-13289
; Sequence 13289, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. HowardTITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13289

LENGTH: 315

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(315)

OTHER INFORMATION: Xaa = Any Amino Acid

US-09-815-242-13289

Query Match

Best Local Similarity 48.7%; Score 38; DB 10; Length 315;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY

1 MRYSALGSDGVRV 15

Db

281 MNYRAKMGFDATRL 295

RESULT 10

US-09-994-064-6

; Sequence 6, Application US/09994064

GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. HowardTITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/994,064

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13289

LENGTH: 315

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(315)

OTHER INFORMATION: Xaa = Any Amino Acid

US-09-815-242-13289

Query Match

Best Local Similarity 48.7%; Score 38; DB 10; Length 315;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY

1 MRYSALGSDGVRV 15

Db

281 MNYRAKMGFDATRL 295

RESULT 11

US-09-994-064-66

; Sequence 66, Application US/09994064

; Patent No. US20030082788A1

GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. HowardTITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/994,064

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13289

LENGTH: 315

TYPE: PRT

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(315)

OTHER INFORMATION: Xaa = Any Amino Acid

US-09-815-242-13289

Query Match

Best Local Similarity 48.7%; Score 38; DB 9; Length 985;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY

3 YRASALGSDGVRVT 16

Db

921 YECTVLISDGRVT 934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/994,064

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/468,190

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 39116-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-994-064-66

Query Match 48.7% Score 38; DB 9; Length 985;
Best Local Similarity 57.1% Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 YRASALGSDGVRV 16
DB 921 YECTVLISDGTREV 934

RESULT 12

US-09-738-626-4312
Sequence 4312, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4312
LENGTH: 1233
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4312

Query Match 48.7% Score 38; DB 9; Length 1233;
Best Local Similarity 37.5% Pred. No. 3.6e+02;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 16
DB 34 IQARISGIGNDGKRT 49

RESULT 13

US-10-156-761-15060
Sequence 15060, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 15060
LENGTH: 355
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-15060

Query Match 48.1% Score 37.5; DB 9; Length 355;
Best Local Similarity 56.2% Pred. No. 1e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 1 MRYRASALGSDGVRV 16
DB 1 MGRFASHS-GIMAT 15

RESULT 14

US-10-156-761-14091
Sequence 14091, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 14091
LENGTH: 172
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-14091

Query Match 47.4% Score 37; DB 9; Length 172;
Best Local Similarity 53.3% Pred. No. 52;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 30 IRYPALLHDDGTRV 44

RESULT 15

US-10-156-761-10243
Sequence 10243, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29

Mon Jul 7 15:30:59 2003

us-09-893-371-2.rapb

Page 6

; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10243
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10243

Query Match 47.4%; Score 37; DB 9; Length 206;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 RASALGSDGVRYT 16
: | | | | : |
Db 119 KAALGSDGVRYT 131

Search completed: July 7, 2003, 14:26:34
Job time : 1.66415 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:12:35 ; Search time 0.815094 Seconds

(without alignments)
1887.083 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78

Sequence: 1 MRYRASALGSDGVRT 16

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR73:
2: PIR:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.8	262	1	enoyl-[acyl]-carrie
2	49	62.8	262	1	enoyl-[acyl]-carrie
3	49	62.8	262	2	enoyl-[acyl]-carrie
4	49	62.8	262	2	enoyl-[acyl]-carrie
5	49	62.8	262	2	enoyl-[acyl]-carrie
6	48	61.5	99	2	enoyl-[acyl]-carrie
7	45	57.7	260	2	enoyl-[acyl]-carrie
8	44	56.4	264	2	enoyl-[acyl]-carrie
9	44	56.4	265	2	enoyl-[acyl]-carrie
10	44	56.4	1112	2	enoyl-[acyl]-carrie
11	44	56.4	1385	2	enoyl-[acyl]-carrie
12	44	56.4	1391	2	enoyl-[acyl]-carrie
13	42	53.8	291	2	enoyl-[acyl]-carrie
14	41	52.6	238	2	enoyl-[acyl]-carrie
15	41	52.6	356	2	enoyl-[acyl]-carrie
16	41	52.6	415	2	enoyl-[acyl]-carrie
17	41	52.6	547	2	enoyl-[acyl]-carrie
18	41	52.6	572	2	enoyl-[acyl]-carrie
19	39	50.0	261	2	enoyl-[acyl]-carrie
20	39	50.0	261	2	enoyl-[acyl]-carrie
21	39	50.0	262	2	enoyl-[acyl]-carrie
22	39	50.0	262	2	enoyl-[acyl]-carrie
23	39	50.0	274	2	enoyl-[acyl]-carrie
24	39	50.0	274	2	enoyl-[acyl]-carrie
25	39	50.0	274	2	enoyl-[acyl]-carrie
26	38.5	49.4	1296	2	enoyl-[acyl]-carrie
27	38.5	49.4	513	2	enoyl-[acyl]-carrie
28	38	48.7	109	2	enoyl-[acyl]-carrie
29	38	48.7	146	2	enoyl-[acyl]-carrie

30	38	48.7	155	2	adenylsulfate 3-
31	38	48.7	272	2	hypothetical prote
32	38	48.7	272	2	enoyl-[acyl]-carrie
33	38	48.7	312	2	AG2669
34	38	48.7	312	2	P95164
35	38	48.7	316	2	Hpr(Ser) kinase/ph
36	38	48.7	374	2	probable hexosyltr
37	38	48.7	378	1	probable flagellar
38	38	48.7	416	2	glutamate-1-semial
39	38	48.7	424	2	hasa export system
40	38	48.7	437	2	conserved hypothet
41	38	48.7	452	2	cyae protein - Bor
42	38	48.7	474	1	hypothetical prote
43	38	48.7	492	2	indole-3-pyruvate
44	38	48.7	513	2	probable sugar tra
45	38	48.7	545	1	
			547	2	

ALIGNMENTS

RESULT 1
S48029
enoyl-[acyl]-carrier-protein] reductase (NADH2) (EC 1.3.1.9) - Escherichia coli (strat
N:Alternate names: enoyl-ACP reductase; short-chain alcohol dehydrogenase homolog env
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S48029; A47681; C64877
R:Kater, M.M.; Koningsstein, G.M.; Nijkamp, H.J.J.; Stultje, A.R.
Plant Mol. Biol. 25, 771-790, 1994
A:Title: The use of a hybrid genetic system to study the functional relationship betw
A:Reference number: S48029; MUID:94355651; PMID:8075395
A:Accession: S48029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <KAT>
A:Cross-references: EMBL:X78733; NID:9587105; PIDN:CA55381.1; PID:9587106
A:Experimental source: strain K-12, substrain W3110
R:Berghaler, H.; Hogenauer, G.; Turnowsky, F.
J. Gen. Microbiol. 138, 2093-2100, 1992
A:Title: Sequences of the env gene and of two mutated alleles in Escherichia coli.
A:Reference number: A47681; MUID:93123967; PMID:1364817
A:Accession: A47681
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <BER>
A:Cross-references: GB:M97219; NID:9145850; PIDN:AA17755.1; PID:9145851
A:Note: sequence extracted from NCBI backbone (NCBI:121825, NCBI:121826)
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A44720; MUID:97426617; PMID:9278503
A:Accession: C64877
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-262 <BIAT>
A:Cross-references: GB:AE000227; GB:U00096; NID:91787543; PIDN:AACT4370.1; PID:917875
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: fabI, envM
C:Function:
A:Description: catalyzes reduction by NADH of enoyl-acyl-carrier-protein to the corre
A:Pathway: fatty acid biosynthesis
A:Note: Inhibited by palmitoyl-CoA and diazaborine
C:Superfamily: enoyl-[acyl]-carrier-protein] reductase (NADH); short-chain alcohol deh
C:Keywords: fatty acid biosynthesis; homotetramer; inner membrane; NAD; oxidoreductas
F:7-190/Domain: short-chain alcohol dehydrogenase homology <SADH>
F:9-35/Region: NAD binding

Query Match 62.8%; Score 49; DB 1; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:|:|:|:|
Db 170 VRYMANAMGPEGV 184

RESULT 2

enoyl-[acyl-carrier-protein] reductase (NADH2) (EC 1.3.1.9) - *Salmonella typhimurium* B43729
C:Species: *Salmonella typhimurium*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: B43729, S70724
R:Turnovsky, F.; Fuchs, K.; Jeschek, C.; Hoegenauer, G.
J. Bacteriol. 171, 6555-6565, 1989
A:Title: envm genes of *Salmonella typhimurium* and *Escherichia coli*.
A:Reference number: A43729; MUID:90078098; PMID:2687243
A:Accession: B43729
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <TUN>
A:Cross-references: GB:M31806; NID:g153953; PIDN:AAA27059.1; PID:g153955
R:Ol, S.Y.; Li, Y.; Sztybel, A.; Gilles, I.G.; Moltr, A.; O'Connor, C.D.
Mol. Microbiol. 17, 523-531, 1995
A:Title: *Salmonella typhimurium* responses to a bactericidal protein from human neutrophils
A:Reference number: S70719; MUID:96100451; PMID:8559071
A:Accession: S70724
A:Molecule type: protein
A:Residues: 7',3'-4',X',6'-11 <QIS>
A:Experimental source: strain SL1344
C:Genetics:
A:Gene: envm
C:Function:
A:Description: catalyzes reduction by NADH of enoyl-acyl-carrier-protein to the correspond
A:Pathway: fatty acid biosynthesis
A>Note: inhibited by diazaboline
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehydro
C:Keywords: fatty acid biosynthesis; inner membrane; NAD; oxidoreductase
F:7-190/Domain: short-chain alcohol dehydrogenase homology <SMAD>
F:9-35/Region: NAD binding

Query Match 62.8%; Score 49; DB 1; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

RESULT 3

enoyl-[acyl-carrier-protein] reductase (NADH) ECS1861 [imported] - *Escherichia coli* (str
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: E90861
R:Hayashi, T.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
Genome Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90861
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <NAV>
A:Cross-references: GB:BA000007; PIDN:BA035284.1; PID:g13361326; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS1861
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehyd

Query Match 62.8%; Score 49; DB 2; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:|:|:|:|
Db 170 VRYMANAMGPEGV 184

RESULT 4

enoyl-[acyl-carrier-protein] reductase (NADH) [imported] - *Escherichia coli* (strain O
H85757
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H85757
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
hiller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <STO>
A:Cross-references: GB:AE005174; NID:g12515500; PIDN:AG56524.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: fadI
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol deh

Query Match 62.8%; Score 49; DB 2; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:|:|:|:|
Db 170 VRYMANAMGPEGV 184

RESULT 5

AD0656
enoyl-[acyl-carrier-protein] reductase (NADH) STY1352 [imported] - *Salmonella enteric*
C:Species: *Salmonella enterica* subsp. *enterica* serovar *typhi*
A>Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD0656
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fair
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0656
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01621.1; PID:g16502475; GSPDB:GN00176
C:Genetics:
A:Gene: STY1352
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol deh

Query Match 62.8%; Score 49; DB 2; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:~|:|:|:|:|
Db 170 VRYMANAMGPEGV 184

RESULT 6

S71883
enoyl-[acyl-carrier-protein] reductase (NADH2) (EC 1.3.1.9) - *Proteus mirabilis* (frag
C:Species: *Proteus mirabilis*
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 03-Jun-2002

C:Accession: S71883
R:Perito, B.; Allocati, N.; Casalone, E.; Masulli, M.; Dragani, B.; Polisinelii, M.; Acet
Biochem. J. 318, 157-162, 1996
A:Title: Molecular cloning and overexpression of a glutathione transferase gene from *Pro*
A:Reference number: S71882; MUID:96358500; PMID:8761466
A:Accession: S71883
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-99 <PER>
A:Cross-references: EMBL:U08482; NID:91053074; PIDN:AAC44361.1; PID:91053075
A:Experimental source: strain AF 2924
C:Function:
A:Description: catalyzes reduction by NADH of enoyl-acyl-carrier-protein to the correspo
A:Pathway: fatty acid biosynthesis
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehyd
C:Keywords: fatty acid biosynthesis; inner membrane; NAD; oxidoreductase

Query Match 61.5%; Score 48; DB 2; Length 99;
Best Local Similarity 53.3%; Pred. No. 0.24;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 7 VRYMANMGPGRIV 21.

RESULT 7
684960
enoyl-[acyl-carrier-protein] reductase (NADH2) (EC 1.3.1.9) [Imported] - Buchnera sp. (S
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 03-Jun-2002
C:Accession: 684960
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: 684960
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: fabI; BU265
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehyd
C:Keywords: oxidoreductase

Query Match 57.7%; Score 45; DB 2; Length 260;
Best Local Similarity 53.3%; Pred. No. 2.3;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 170 VRYMASLCKENIRV 184

RESULT 8
AG3354
enoyl-[acyl-carrier-protein] reductase [Imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AG3354
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium *Ana*
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG3354
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA876090.1; PID:917133527; GSPDB:GN00179

A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: a114391
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehy

Query Match 56.4%; Score 44; DB 2; Length 264;
Best Local Similarity 60.0%; Pred. No. 3.5;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 178 VRYLASLGSQNTIRV 192

RESULT 9
C83419
NADH-dependent enoyl-ACP reductase PA1806 [Imported] - *Pseudomonas aeruginosa* (strain
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83419
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardi, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83419
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <STO>
A:Cross-references: GB:AE004607; GB:AE004091; NID:99947789; PIDN:AAG05195.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: fabI; PA1806
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehy

Query Match 56.4%; Score 44; DB 2; Length 265;
Best Local Similarity 53.3%; Pred. No. 3.5;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 173 VRYLAGSLGAEGRIV 187

RESULT 10
S28289
hypothetical protein C38C10.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jan-2000
C:Accession: S28289
R:Thoms, K.
submitted to the EMBL Data Library, December 1992
A:Reference number: S28285
A:Accession: S28289
A:Molecule type: DNA
A:Residues: 1-1112 <THO>
A:Cross-references: EMBL:219153
C:Genetics:
A:Insertions: 412/1; 612/2; 670/3; 676/3; 729/3; 914/2; 984/1
C:Superfamily: *Caenorhabditis elegans* hypothetical protein C38C10.5

Query Match 56.4%; Score 44; DB 2; Length 1112;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 16
DB 1040 IRRASOMNGDVNAT 1055

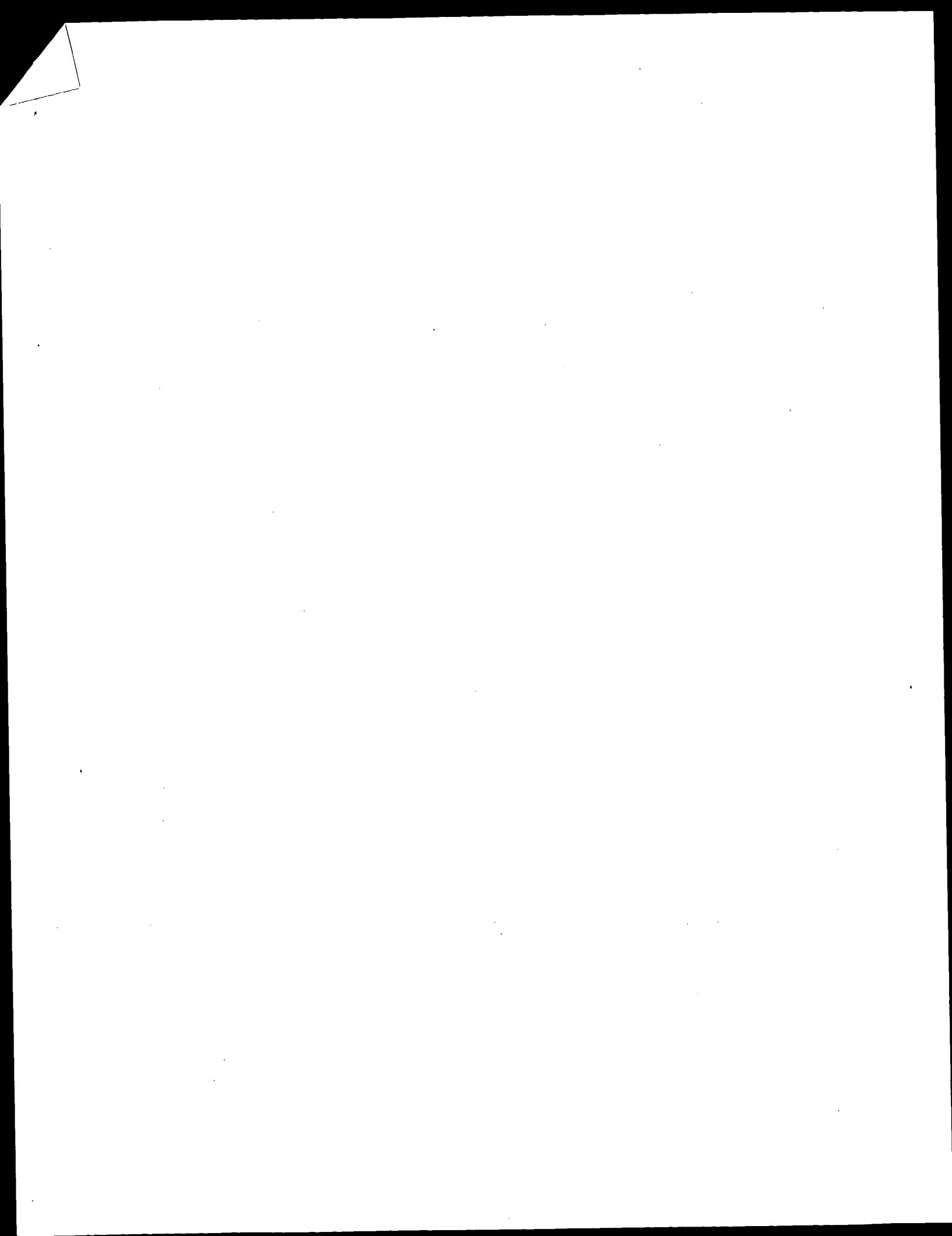
RESULT 11
AB8554

A:Accession: T04665
A:Molecule type: DNA
A:Residues: 1-356 <BEV>
A:Cross-References: EMBL:AL031135; GSPDB:GN00062; ATSP:F8D20.110
A:Experimental source: cultivar Columbia; BAC clone F8D20
C:Genetics:
A:Gene: ATSP:F8D20.110
A:Map position: 4
A:Note: Intron positions not resolved
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: phosphotransferase; protein kinase
F:21-309/Domain: protein kinase homology <KIN>

Query Match 52.6%; Score 41; DB 2; Length 356;
Best Local Similarity 53.3%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRYT 16
DB 333 RYKAGALGAERKRAT 347

Search completed: July 7, 2003, 14:24:55
Job time : 1.81509 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:44 ; Search time 0.45283 Seconds

(without alignments)
1465.497 Million cell updates/sec

Title: US-09-893-371-2

Sequence: 1 MKRASALGSDGVRVT 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	62.8	261	1	FABI_ECOLI
2	49	62.8	261	1	FABI_SALTY
3	45	57.7	260	1	FABI_BUCAT
4	44	56.4	265	1	FABI_ANASP
5	44	56.4	265	1	FABI_PSEAE
6	44	56.4	1391	1	YLD5_CAEEL
7	41	52.6	548	1	AMT4_PSEST
8	39	50.0	268	1	FAI2_RHIME
9	38	48.7	109	1	RLAI_MAIZE
10	38	48.7	155	1	CYSC_ARCFU
11	38	48.7	229	1	TPIS_SULTO
12	38	48.7	272	1	FAI1_RHIME
13	38	48.7	424	1	GSA_CAMTE
14	38	48.7	474	1	CYAE_BORE
15	38	48.7	513	1	YH26_MZCVR
16	38	48.7	545	1	DCIP_AZOB
17	37	47.4	353	1	H182_MYCTU
18	37	47.4	551	1	AMT4_PSESA
19	37	47.4	622	1	CZAC_BACRU
20	37	47.4	2212	1	RRPL_EROCM
21	36	46.2	229	1	HIS4_PYRAE
22	36	46.2	237	1	YFV4_METTF
23	36	46.2	258	1	FABI_SYNT3
24	36	46.2	261	1	FABI_HAEIN
25	36	46.2	334	1	BCHI_RHOSE
26	36	46.2	350	1	BCHI_RHOSE
27	36	46.2	361	1	RPL_MTCLE
28	36	46.2	401	1	DKR_STRCO
29	36	46.2	427	1	TOLE_HAEIN
30	36	46.2	617	1	ESRI_ICFPU
31	36	46.2	710	1	DAS_PICAN
32	36	46.2	715	1	AT12_HSVIF
33	36	46.2	718	1	AT12_HSV11

34	36	46.2	778	1	HTR6_HAIUN	Q48392 halobacteri
35	36	46.2	778	1	HTR6_HAIUN	Q48392 halobacteri
36	36	46.2	817	1	PPSA_PYRFB	P42850 pyrococcus
37	36	46.2	819	1	PPSA_PYRFB	O92877 pyrococcus
38	36	46.2	821	1	PPSA_PYRFB	O57830 pyrococcus
39	36	46.2	1004	1	SIPO_BACER	P09333 bacillus br
40	36	46.2	1188	1	PPSA_METUA	O57962 methanococc
41	35	44.9	258	1	FABI_BACSU	P54616 bacillus su
42	35	44.9	285	1	G539_BACSU	P80873 bacillus su
43	35	44.9	353	1	VGLM_MCMWK	P52373 murine cyto
44	35	44.9	360	1	MTDH_ARATH	P42734 arabidopsis
45	35	44.9	368	1	RP2_STRCO	O53915 streptomyces

ALIGNMENTS

RESULT 1	ID	FABI_ECOLI	STANDARD	PRT	261 AA
AC	P29132				
DT	01-DEC-1992 (Rel. 24, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-dependent enoyl-ACP reductase)				
GN	FABI OR ENVW OR B1268 OR Z2512 OR ECS1861.				
OS	Escherichia coli, and				
OC	Escherichia coli O157:H7.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.				
OX	NCBI_TaxID=562, 83334;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-30.				
RX	MEDLINE=93123967; PubMed=1364817;				
RA	Bergler H., Hoegenauer G., Turnowsky F.;				
RT	"Sequences of the envm gene and of two mutated alleles in Escherichia coli".				
RL	J. Gen. Microbiol. 138:2093-2100(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-K12 / W3110;				
RX	MEDLINE=94355651; PubMed=8075395;				
RA	Kater M.M., Konigstein G.M., Nijkamp H.J.J., Stultje A.R.;				
RT	"The use of a hybrid genetic system to study the functional relationship between prokaryotic and plant multi-enzyme fatty acid synthetase complexes".				
RL	Plant Mol. Biol. 25:771-790(1994).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-K12 / MG1655.				
RX	MEDLINE=97426617; PubMed=9278503;				
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,				
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,				
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,				
RA	Mau B., Shao Y.;				
RT	"The complete genome sequence of Escherichia coli K-12".				
RL	Science 277:1453-1474(1997).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-K12;				
RX	MEDLINE=97251357; PubMed=9097039;				
RA	Alba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,				
RA	Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,				
RA	Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,				
RA	Nakase S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,				
RA	Takemoto K., Seki Y., Sivasubramanian S., Tagami H., Takeda J.,				
RA	Sampel G., Takeuchi Y., Wada C., Yamamoto Y., Horikuchi T.;				
RT	"A 570-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map".				
RL	DNA Res. 3:363-377(1996).				
RN	[5]				
RP	SEQUENCE FROM N.A.				

RC STRAIN-0157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE-21074935; PubMed-11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Mantharman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RL "Genome sequence of enterohaemorrhagic *Escherichia coli* 0157:H7,"
 Nature 409:529-533(2001).
 [6]
 RN SEQUENCE FROM N.A.
 RP STRAIN-0157:H7 / RIND 0509952;
 RX MEDLINE-21156231; PubMed-11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RL "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 0157:H7 and genomic comparison with a laboratory strain K-12,"
 DNA Res. 8:11-22(2001).
 [7]
 RN PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RP MEDLINE-94164884; PubMed-8119879;
 RX Berger H., Wallner P., Ebeling A., Leitinger B., Fuchsichler S.,
 RA Aschner H., Kollenz G., Hoegenauer G., Turnowsky F.;
 RL "Protein Envm is the NADH-dependent enoyl-ACP reductase (FabI) of
Escherichia coli,"
 J. Biol. Chem. 269:5493-5496(1994).
 [8]
 RN SEQUENCE OF 1-12.
 RP STRAIN-K12 / EMG2;
 RX MEDLINE-97443975; PubMed-9298646;
 RA Link A.J., Robinson K., Church G.M.;
 RL "Comparing the predicted and observed properties of proteins encoded
 in the genome of *Escherichia coli* K-12,"
 Electrophoresis 18:1259-1313(1997).
 [9]
 RN SEQUENCE OF 1-11.
 RP STRAIN-K12 / W3110;
 RX Frutiger S., Hughes G.J., Pasquali C., Hochstrasser D.F.;
 RL Submitted (FEB-1996) to the SWISS-PROT data bank.
 [10]
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RP MEDLINE-97113207; PubMed-8953047;
 RX Baldoek C., Rafferty J.B., Sedelnikova S.E., Baker P.J., Stuitje A.R.,
 RA Slabas A.R., Hawkes P.R., Rice D.W.;
 RL "A mechanism of drug action revealed by structural studies of enoyl
 reductase,"
 Science 274:2107-2110(1996).
 [11]
 RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RP MEDLINE-99425142; PubMed-10493822;
 RX Ward W.H., Holdgate G.A., Rowsell S., McLean E.G., Pauplit R.A.,
 RA Clayton E., Nichols W.W., Collis J.G., Minshull C.A., Jude D.A.,
 RA Mistry A., Timms D., Camble R., Hales N.J., Brilton C.J.,
 RA Taylor I.W.;
 RL "Kinetic and structural characteristics of the inhibition of enoyl
 (acyl carrier protein) reductase by triclosan,"
 Biochemistry 38:12514-12525(1999).
 [12]
 RN X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RP MEDLINE-99329134; PubMed-10398587;
 RX Stewart M.J., Parikh S., Xiao G., Tonge P.J., Kisker C.;
 RL "Structural basis and mechanism of enoyl reductase inhibition by
 triclosan,"
 J. Mol. Biol. 290:859-865(1999).
 [13]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RP MEDLINE-99215552; PubMed-10201369;
 RX Levy C.W., Roujeinikova A., Sedelnikova S., Baker P.J., Stuitje A.R.,
 RA Slabas A.R., Rice D.W., Rafferty J.B.;
 RL "Molecular basis of triclosan activity,"

RL Nature 398:383-384(1999).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) - trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated.
 CC -1- MISCELLANEOUS: THE ANTIBIOTIC DIAZABORINE INTERFERES WITH THE
 CC ACTIVITY BY BINDING TO THE PROTEIN.
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M97219; AAA17755.1; -;
 CC EMBL: X78733; CA55381.1; -;
 CC EMBL: AE000227; AAC74370.1; -;
 CC EMBL: D90767; BA114841.1; -;
 CC EMBL: D90767; BA114849.1; -;
 CC EMBL: AE005378; AAC56524.1; -;
 CC EMBL: AP002556; BAB35284.1; -;
 CC PIR: A47681; A47681.
 CC PDB: 1DFG; 28-JAN-98.
 CC PDB: 1DFH; 28-JAN-98.
 CC PDB: 1DFI; 28-JAN-98.
 CC PDB: 1QG6; 21-SEP-99.
 CC PDB: 1QSG; 21-JUL-99.
 CC PDB: 1DBA; 28-OCT-99.
 CC SWISS-2DPAGE: P29132; COLI.
 CC Ecocore: BG11528; fabI.
 CC InterPro: IPR002198; ADH_short.
 CC Pfam: PF00106; adh_short; 1.
 CC Oxidoreductase; NAD; Fatty acid biosynthesis; Antibiotic resistance;
 CC Inner membrane; 3D-structure; Complete proteome.
 CC INIT MET 0 0
 CC NP_BIND 9 35 NAD (BY SIMILARITY).
 CC MUTAGEN 92 92 G->S: DIAZABORINE RESISTANCE.
 CC MUTAGEN 240 240 S->R: PRODUCES TEMPERATURE-SENSITIVE
 CC PHENOTYPE.
 CC SQ SEQUENCE 261 AA; 27732 MW; 434B019F34855956 CRC64;
 CC Query Match 62.8%; Score 49; DB 1; Length 261;
 CC Best Local Similarity 60.0%; Pred. No. 0.29;
 CC Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 1 MRYASALGSDGVY 15
 CC :||:|:|:|
 CC Db 169 VRYMANMGPQGVY 183
 CC
 CC RESULT 2
 CC ID FABI_SALTY STANDARD; PRT; 261 AA.
 CC AC P16657;
 CC DT 01-AUG-1990 (Rel. 15, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
 CC dependent enoyl-ACP reductase).
 CC GN FABI OR ENVM OR STM1700.
 CC OS Salmonella typhimurium.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC OC Salmonella.
 CC OX NCBI_TaxID=602;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-AG701;
 CC RX MEDLINE-90078098; PubMed-2687243;

RA Turnowsky F., Fuchs K., Jeschek C., Hoegenauer G.;
 RT "envm genes of *Salmonella typhimurium* and *Escherichia coli*.";
 RL Bacteriol. 171:6555-6565(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-L72 / SGGC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Speth J., Clifton S.W., Lattelle P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RL L72.";
 CC Nature 413:852-856(2001).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated.
 CC -1- MISCELLANEOUS: THE ANTIBIOTIC DIAZABORINE INTERFERES WITH THE
 CC ACTIVITY BY BINDING TO THE PROTEIN.
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M31806; AAA27059.1; -
 DR EMBL: AE008775; AAL20618.1; -
 DR PIR: B43729; B43729.
 DR HSSP: P29132; 1DPT.
 DR Styrene; SG10035; fadi.
 DR Interpro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Oxidoreductase; NAD; Fatty acid biosynthesis; Antibiotic resistance;
 KW Inner membrane; Complete proteome.
 FT INIT_MET 0 BY SIMILARITY.
 FT NP_BIND 9 35 NAD (BY SIMILARITY).
 FT MUTAGEN 92 92 G->S: DIAZABORINE RESISTANCE.
 FT SEQUENCE 261 AA; 27629 MW; 187B91AE341B773 CRC64;
 SQ
 Query Match 62.8%; Score 49; DB 1; Length 261;
 Best Local Similarity 60.0%; Pred. No. 0.29;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVR 15
 DB 169 VRYMANMGPCGVR 183
 RESULT 3
 FABI_BUCAT STANDARD: PRT; 260 AA.
 AC P57353;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
 DE dependent enoyl-ACP reductase).
 GN FABI OR B0265.
 OS *Buchnera aphidicola* (subsp. *Acyrthosiphon pisum*) (*Acyrthosiphon pisum*
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; *Buchnera*.
 CC NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Tokyo 1998;
 RX MEDLINE-20445173; PubMed-10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RL *Buchnera* sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADH.
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AP001118; BAB12975.1; -
 DR HSSP: P29132; 1DPT.
 KW Oxidoreductase; NAD; Fatty acid biosynthesis; Complete proteome.
 FT NP_BIND 10 36 NAD (BY SIMILARITY).
 FT SEQUENCE 260 AA; 28779 MW; DCA04AC740D6DADD CRC64;
 SQ
 Query Match 57.7%; Score 45; DB 1; Length 260;
 Best Local Similarity 53.3%; Pred. No. 1.4;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVR 15
 DB 170 VRYMANMGPCGVR 184
 RESULT 4
 FABI_ANASP STANDARD: PRT; 258 AA.
 AC Q05069;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
 DE dependent enoyl-ACP reductase).
 GN FABI OR AL4391.
 OS *Anabaena* sp. (strain PCC 7120).
 CC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 CC NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-93308081; PubMed-8391534;
 RA Wei T.-F., Ramasubramanian T.S., Pu F., Golden J.W.;
 RT "Anabaena sp. strain PCC 7120 b1fa gene encoding a sequence-specific
 RT DNA-binding protein cloned by in vivo transcriptional interference
 RL selection.";
 RL J. Bacteriol. 175:4025-4035(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-21595285; PubMed-11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida T., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium *Anabaena* sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L10036; AAD04184.1; ALT_INIT.
 DR EMBL: AP003596; BAB76090.1; ALT_INIT.
 DR HSSP: P29132; IDFI.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KM Oxidoreductase; NAD; Fatty acid biosynthesis; Complete proteome.
 FT NP_BIND 10 36 NAD (BY SIMILARITY).
 SQ SEQUENCE 258 AA; 27421 MW; 45012CA8497E001B CRC64;

Query Match 56.4%; Score 44; DB 1; Length 258;
 Best Local Similarity 60.0%; Pred. No. 2.1;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 MRYRASALGSDGVRV 15
 Db 172 VRYLASELGSOMIRV 186

RESULT 5
 FABI_PSEAE STANDARD; PRT; 265 AA.

AC Q9ZFE4;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
 DE dependent enoyl-ACP reductase).
 OS FABI OR PA1806.
 GN Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NC NCBL_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=99395061; PubMed=10464225;
 RA Hoang T.T., Schweizer H.P.;
 RT "Characterization of Pseudomonas aeruginosa enoyl-acyl carrier protein
 RT reductase (Fabi): a target for the antimicrobial triclosan and its
 RT role in acylated homoserine lactone synthesis.";
 RL J. Bacteriol. 181:5489-5497(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltz L., Tolentino E., Westbrook-Wadman S., Yan Y.,
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Laible K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: AF104262; AAC95362.1;
 DR EMBL: AE004607; AAG05195.1;
 DR HSSP: P29132; IDFI.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KM Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;
 FT NP_BIND 10 36 NAD (BY SIMILARITY).
 SQ SEQUENCE 265 AA; 28006 MW; 3E7A946886DFB293 CRC64;

Query Match 56.4%; Score 44; DB 1; Length 265;
 Best Local Similarity 53.3%; Pred. No. 2.2;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 MRYRASALGSDGVRV 15
 Db 173 VRYLAGSLGAGTRV 187

RESULT 6
 YLD5_CAEEL STANDARD; PRT; 1391 AA.

AC Q03570; P39218; P34581;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C38C10.5 in chromosome III.
 GN C38C10.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NC NCBL_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Cosey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnstone L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lighting J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Woldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP REVISIONS, AND ALTERNATIVE SPLICING.
 RA Jones S.J.M.;
 RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a and b (shown here); may be
 CC produced by alternative splicing.
 CC -1- SIMILARITY: TO YEAST RGR1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z29115; CAAB2365.1; JOINED.
 DR EMBL: Z19153; CAAB2365.1; JOINED.
 DR EMBL: Z19153; CAAB2365.1; JOINED.
 DR EMBL: Z29115; CAAB2366.1; JOINED.
 DR EMBL: Z29115; CAAB2366.1; JOINED.

DR EMBL; 219153; CAA82366.1; JOINED.
 DR EMBL; 219153; CAA79551.1; -
 DR EMBL; 229115; CAA79551.1; JOINED.
 DR PIR; S28289; S28289; JOINED.
 DR WormRep; C38C10.5; CE20550.
 KW Hypothetical protein; Alternative splicing.
 FT VARSPLIC 671 676 MISSING (IN ISOFORM A).
 SQ SEQUENCE 1391 AA; 157165 MW; 8702092DAC55A4BF CRC64;
 Query Match 56.48; Score 44; DB 1; Length 1391;
 Best Local Similarity 50.08; Pred. No. 13;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MRYRASALGSDGVAVT 16
 Db 1040 IRRFASQMGDGVAVT 1055
 RESULT 7
 ID AMT4_PSEST STANDARD; PRT; 548 AA.
 AC P13507;
 DT 01-JUN-1990 (Rel. 13, Created)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glucan 1,4-alpha-maltotetrahydrolase precursor (EC 3.2.1.60) (G4-
 DE amylose) (Maltotetraose-forming amylase) (Exo-maltotetrahydrolase)
 DE (Maltotetraose-forming exo-amylase).
 GN AMYP.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-MO-19;
 RX MEDLINE=89155431; PubMed=2646279;
 RA Fujita M., Torigoe K., Nakada T., Tsusaki K., Kubota M., Sakai S.,
 RA Tsujisaka Y.;
 RT "Cloning and nucleotide sequence of the gene (amyP) for
 RT maltotetraose-forming amylase from Pseudomonas stutzeri MO-19.";
 RL J. Bacteriol. 171:1333-1339(1989).
 [2]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND REVISIONS TO 286-302.
 RC STRAIN-MO-19;
 RX MEDLINE=97271999; PubMed=9126844;
 RA Morishita Y., Hasegawa K., Matsura Y., Katsube Y., Kubota M.,
 RA Sakai S.;
 RT "Crystal structure of a maltotetraose-forming exo-amylase from
 RT Pseudomonas stutzeri.";
 RL J. Mol. Biol. 267:661-672(1997).
 [3]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF VARIANT GLN-240.
 RC STRAIN-MO-19;
 RX MEDLINE=97428332; PubMed=9281429;
 RA Yoshioka Y., Hasegawa K., Matsura Y., Katsube Y., Kubota M.;
 RT "Crystal structures of a mutant maltotetraose-forming exo-amylase
 RT co-crystallized with maltopentaose.";
 RL J. Mol. Biol. 271:619-628(1997).
 [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF VARIANTS.
 RC STRAIN-MO-19;
 RX MEDLINE=20027472; PubMed=10556241;
 RA Hasegawa K., Kubota M., Matsura Y.;
 RT "Roles of catalytic residues in alpha-amylases as evidenced by the
 RT structures of the product-complexed mutants of a maltotetraose-forming
 RT amylase.";
 RL Protein Eng. 12:819-824(1999).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-alpha-D-glucosidic linkages
 CC in amyloseous polysaccharides so as to remove successive
 CC maltotetraose residues from the non-reducing chain ends.
 CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
 CC -1- PATHWAY: Starch degradation.

CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVERAL ISOENZYME FORMS OF THIS PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 DR EMBL; M24516; AAA25707.1; -
 DR PIR; A32803; A32803.
 DR PDB; 2AMG; 01-APR-97.
 DR PDB; 1JDA; 15-OCT-97.
 DR PDB; 1JDC; 15-OCT-97.
 DR PDB; 1JDD; 15-OCT-97.
 DR PDB; 1Q13; 24-NOV-99.
 DR PDB; 1Q14; 24-NOV-99.
 DR PDB; 1Q15; 24-NOV-99.
 DR PDB; 1QPK; 17-NOV-99.
 DR InterPro: IPR000461; Alpha_amylase.
 DR InterPro: IPR002044; CBD_4.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF00686; CBM_20; 1.
 DR ProDom: PD001568; CBD_4; 1.
 KW Hydrolyase; Glycosidase; Signal; Carbohydrate metabolism; Calcium;
 KW 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 1 548
 FT DISULFID 161 171
 FT DISULFID 237 272
 FT CA_BIND 22 22
 FT CA_BIND 23 23
 FT CA_BIND 23 23
 FT CA_BIND 34 34
 FT CA_BIND 37 37
 FT CA_BIND 38 38
 FT CA_BIND 137 137
 FT CA_BIND 172 172
 FT CA_BIND 175 175
 FT CA_BIND 183 183
 FT ACT_SITE 214 214
 FT ACT_SITE 240 240
 FT ACT_SITE 315 315
 FT ACT_SITE 315 315
 FT CONFLICT 286 302
 SQ SEQUENCE 548 AA; 59876 MW; 2887217B379158F CRC64;
 REF. 1).
 GSIAWMKHLGNGNDPDR -> ARSPTGSTERQSRPA (IN
 Query Match 52.68; Score 41; DB 1; Length 548;
 Best Local Similarity 60.08; Pred. No. 16;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MRYRASALGSDGVAVT 15
 Db 116 LRQAASALGAGVAVT 130
 RESULT 8
 ID FA12_RHIME STANDARD; PRT; 268 AA.
 AC P58381;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enoyl-[acyl-carrier-protein] reductase [NADH] 2 (EC 1.3.1.9) (NADH-
 DE dependent enoyl-ACP reductase 2).
 GN FAB12 OR R00246 OR SMC00326.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobium.
 OX NCBI_TaxID=382;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Batloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 Bostard P., Becker A., Boutry M., Cadieu E., Drenno S., Gloux S.,
 Godie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
 Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 Rhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -1- CATALYTIC ACTIVITY: Acyl-(acyl-carrier protein) + NAD(+) = trans-
 2,3-dehydroacyl-(acyl-carrier protein) + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (by similarity).
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL591783; CAC41683.1; -;
 CC InterPro: IPR002198; ADH_short.
 CC Pfam: PF00106; adh_short.1.
 CC Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;
 KM Complete proteome.
 FT NP_BIND 11 37 NAD (BY SIMILARITY).
 SQ SEQUENCE 268 AA; 28671 MW; 83D2127C7D76628C CRC64;
 QY Query Match 50.0%; Score 39; DB 1; Length 268;
 Best Local Similarity 53.3%; Pred. No. 16;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 DB 1 MRYRASALGSDGVRV 15
 170 VRYLADYSGRGIRV 184
 DB
 RESULT 9
 RLAL_MAIZE
 ID RLAL_MAIZE STANDARD; PRT; 109 AA.
 AC P52855; O24414;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 60S acidic ribosomal protein P1 (L12).
 GN RPL1A.
 OS Zea mays (Maize).
 CC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC Clade;
 CC Panicoideae; Andropogoneae; Zea.
 CC NCBI_TaxID=4577;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. W22;
 RC Hamilton D.A., Turcich M.P., Bokhari-Riza A., Mascarenhas J.P.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 18-21.
 RC STRAIN=cv. B73; TISSUE=ear;
 RX MBLDLINE=97422884; PubMed=9276949;
 RA Bailey-Serres J., Vangala S., Szick K., Lee C.H.;
 RT "Acidic phosphoprotein complex of the 60S ribosomal subunit of maize
 seedling roots. Components and changes in response to flooding.";
 RL Plant Physiol. 114:1293-1305(1997).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
 CC PROTEIN SYNTHESIS (BY SIMILARITY).

CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
 CC SUBUNIT (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED.
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U040147; AAA91168.1; -;
 CC EMBL: U62752; AAB71079.1; -;
 CC DR MaltZDB: 84941; -;
 CC DR InterPro: IPR001813; 60S-ribosomal.
 CC DR Pfam: PF00428; 60S-ribosomal; 1.
 CC KW Ribosomal protein.
 CC CONFLICT 9
 FT SEQUENCE 109 AA; 11096 MW; 1BFEC8E34391F080 CRC64;
 SQ
 QY Query Match 48.7%; Score 38; DB 1; Length 109;
 Best Local Similarity 46.7%; Pred. No. 9.1;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 DB 2 RYRASALGSDGVRV 16
 9 RYALILSDSGIAIT 23
 DB
 RESULT 10
 CYSC_ARCFU
 ID CYSC_ARCFU STANDARD; PRT; 155 AA.
 AC Q29533;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
 DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
 DE phosphotransferase).
 GN CYSC OR AF0286.
 OS Archaeoglobus fulgidus.
 CC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 CC Archaeoglobaceae; Archaeoglobus.
 CC NCBI_TaxID=2234;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RC MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerkvliet A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirnesh E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reisch C.I., McNeill L.K., Badger J.H., Glodet A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
 CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
 CC phosphoadenylylsulfate.
 CC -1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. INVOLVED
 CC IN THE ENERGY-GENERATING SULFUR OXIDATION PATHWAYS OF MANY
 CC CHEMOAUTOTROPHIC MICROBS.
 CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE001085; AAB90945.1; -

DR TIGR: AF0288; -

DR InterPro: IPR002891; APS_kinase.

DR Pfam: PF01583; APS_kinase; 1.

DR ProDom: PD002350; APS_kinase; 1.

DR TIGRfams: TIGR00455; apsk; 1.

DR Transferase: Kinase; ATP-binding; Phosphorylation; Complete proteome.

FT NP_BIND 9 16 ATP (BY SIMILARITY).

FT ACT_SITE 83 83 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY

SO SEQUENCE 155 AA; 17883 MW; 4535298F48931881 CRC64;

Query Match 48.7%; Score 38; DB 1; Length 155;

Best Local Similarity 57.1%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1 MRYRASALGSDGVR 14

29 MGYRVELDGDGVR 42

RESULT 11

TPIS_SULTO STANDARD; PRT; 229 AA.

AC G96Y29;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Triosephosphate isomerase (EC 5.3.1.1) (TIM).

OS Sulfolobus tokodaii.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_TaxID=111955;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JCM 10545 / 7;

RA MEDLINE=21456156; PubMed=11572479;

RA Kawaiyabashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,

RA Sekine M., Baba S.-I., Ankel A., Kosugi H., Hosoyama A., Fukui S.,

RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,

RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,

RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,

RA Oshima T., Kikuchi H.;

RT "Complete genome sequence of an aerobic thermoacidophilic

RT Crenarchaeon, Sulfolobus tokodaii strain7."

RL DNA Res. 8:123-140(2001).

CC -1 CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate - glycercene

CC phosphate.

CC -1 PATHWAY: Plays an important role in several metabolic pathways.

CC -1 SUBUNIT: Homotrimer (By similarity).

CC -1 SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AP000988; BAB67127.1; -

DR InterPro: IPR003009; FMN_enzyme.

DR InterPro: IPR000652; Triophos_ismrse.

DR ProDom: PD001005; Triophos_ismrse; 1.

DR TIGRfams: TIGR00419; tim; 1.

DR PROSITE: PS00171; TIM; FALSE NEG.

KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;

KW Pentose shunt; Complete proteome.

FT ACT_SITE 93 93 BY SIMILARITY.

FT SEQUENCE 229 AA; 24898 MW; 9DC1D550A0183525 CRC64;

Query Match 48.7%; Score 38; DB 1; Length 229;

Best Local Similarity 53.8%; Pred. No. 20;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 3 YRASALGSDGVR 15

188 YKAEIGADGIGV 200

RESULT 12

FALL_RHIME STANDARD; PRT; 272 AA.

AC P58360;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Enoyl-lacyl-carrier-protein reductase [NADH] 1 (EC 1.3.1.9) (NADH-

DE dependent enoyl-ACP reductase 1).

GN FABI1 OR R00898 OR SMC00005.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RA MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,

RA Pohl T., Portelle D., Puehler A., Fumelle B., Rampeger U.,

RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021."

RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

CC -1 CATALYTIC ACTIVITY: Acyl-lacyl-carrier protein] + NADH -

CC 2,3-dehydroacyl-lacyl-carrier protein] + NADH.

CC -1 PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.

CC -1 SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).

CC -1 SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN

CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AL591785; CAC45470.1; -

DR InterPro: IPR002198; ADH_short.

DR Pfam: PF00106; adh_short; 1.

KW Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;

KW Complete proteome.

FT NP_BIND 14 40 NAD (BY SIMILARITY).

SO SEQUENCE 272 AA; 29148 MW; 1E0FA18A22CDBE36 CRC64;

Query Match 48.7%; Score 38; DB 1; Length 272;

Best Local Similarity 53.3%; Pred. No. 24;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 1 MRYRASALGSDGVR 15

173 VRYLAVDGNRGIRV 187

```

RESULT 13
GSA_CAMJF
ID GSA_CAMJF STANDARD: PRT: 424 AA.
AC 09PT0:
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) (GSA)
DE (Glutamate-1-semialdehyde aminotransferase) (GSA-AT).
GN HEML OR CJO853.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
RN NCBI_TaxID=197;
RP [1]
SEQUENCE FROM N.A.
RC STRAIN=NCCT 1168; PubMed-10688204;
RA Parhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S.,
RA Jags K., Rajshree A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Randle M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
CC -1- CATALYTIC ACTIVITY: (S)-4-amino-5-oxopentanoate - 5-
CC aminolevulinate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; second step.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL139076; CAB73118.1; -.
DR HSSP: P24630; 2GSA.
DR InterPro: IPR000954; Aminoctran_3.
DR Pfam: PF00202; aminoctran_3; 1.
DR TRIGRAMS: TRIGR00713; heml; 1.
DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; FALSE_NEG.
KM Porphyrin biosynthesis; Isomerase; Pyridoxal phosphate;
KM Complete proteome. 263 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT BINDING 263
FT SEQUENCE 424 AA; 46092 MW; 781152645ADAA57D CRC64;
SQ
Query Match 48.7%; Score 38; DB 1; Length 424;
Best Local Similarity 72.7%; Pred. No. 39;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 YRASALGSGV 13
DB 241 YRASALGSGYI 251

```

```

RESULT 14
CYAE_BORPE STANDARD: PRT: 474 AA.
ID CYAE_BORPE
AC P11092;
DR 01-JUL-1989 (Rel. 11, Created)
DR 01-JUL-1989 (Rel. 11, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein cyae precursor.
RN

```

```

GN CYAE.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18323;
RC MEDLINE=89091151; PubMed-2905265;
RA Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;
RT "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-
RT haemolysin bifunctional protein of Bordetella pertussis."
RL EMBO J. 7:3997-4004(1988).
CC -1- FUNCTION: CYAE IS NECESSARY FOR TRANSPORT OF CALMODULIN-SENSITIVE
CC ADENYLATE CYCLASE-HEMOLYSIN (CYCLOLYSIN).
CC -1- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PRTF FAMILY OF SECRETION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14199; CAA32414.1; -.
DR PIR: S02388; BVRBCE.
DR InterPro: IPR003423; OEP.
DR Pfam: PF02321; OEP; 2.
KM Hemolysis; Transport; Outer membrane; Signal.
FT SIGNAL 1 31
FT CHAIN 32 474 POTENTIAL.
FT PROTEIN CYAE.
SQ SEQUENCE 474 AA; 50204 MW; 29A4F21B377FC957 CRC64;
QY 4 RASALGSGV 14
DB 115 RSTGLGEGDVR 125
Query Match 48.7%; Score 38; DB 1; Length 474;
Best Local Similarity 63.6%; Pred. No. 44;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 RASALGSGV 14
DB 115 RSTGLGEGDVR 125

```

```

RESULT 15
Y26_MYCTU STANDARD: PRT: 513 AA.
ID Y26_MYCTU
AC Q10510;
DR 01-OCT-1996 (Rel. 34, Created)
DR 01-OCT-1996 (Rel. 34, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2226.
GN RV2226 OR MT2285 OR MTCY427.07.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RC MEDLINE=98295987; PubMed-9634220;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jags K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]

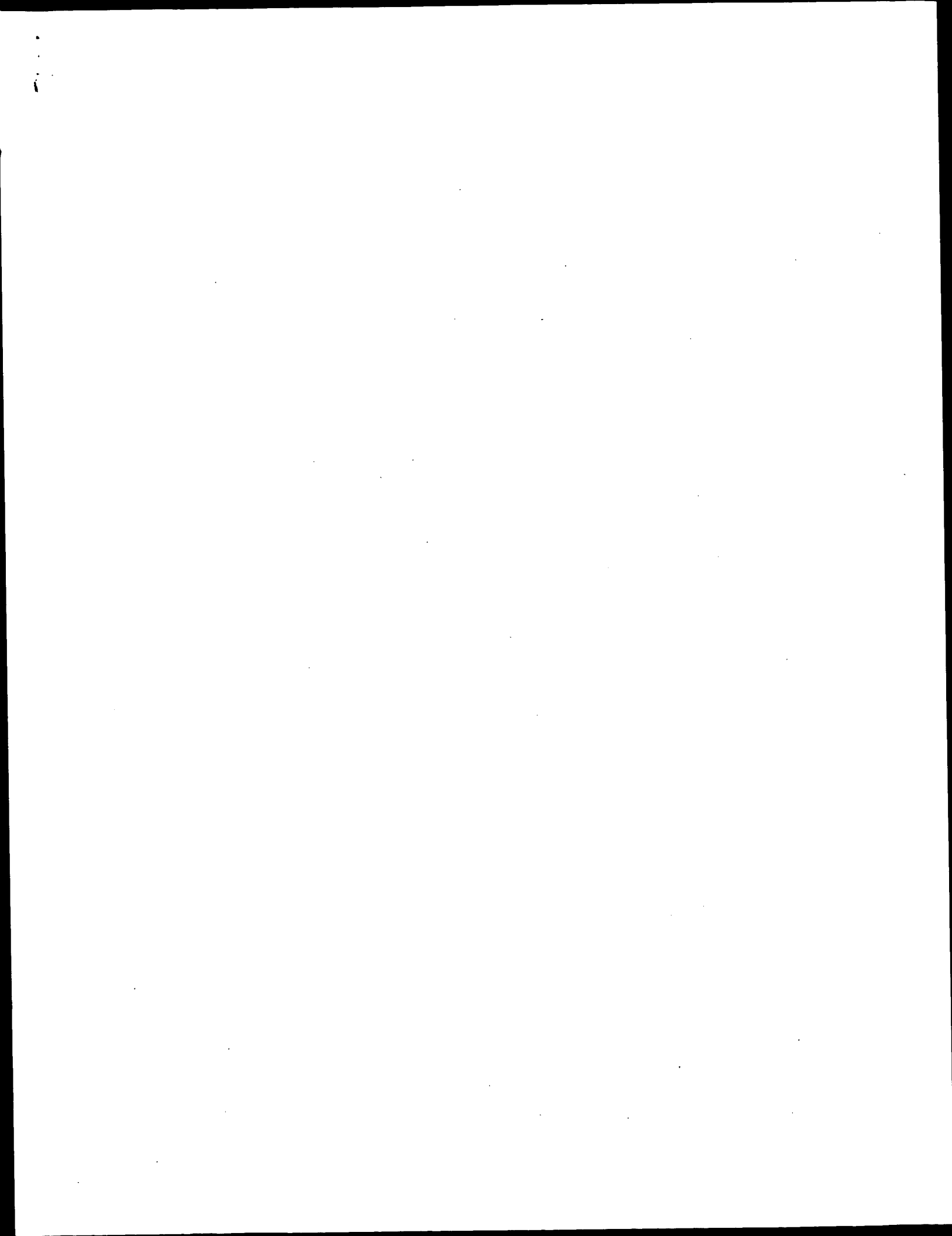
```

RP SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh:
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Debay R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA DeCher A., Utterback T., Weidman J., Khouri H., Gail J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z70692; CAA94649.1; -
DR EMBL; AE007073; AAK46570.1; -
DR TIGR; MT2285; -
DR Tuberculist; RV2226; -
KM Hypothetical protein; Complete proteome.
FT CONFLICT 299 299 D -> N (IN REF. 2).
SO SEQUENCE 513 AA; 56333 MW; 8E74BDA8945D7B99 CRC64;

QY 1 MRRASALGSDGV 13
: 1 1 1 1 1 1 1
Db 426 LRYTAATGADNV 438

Query Match 48.7%; Score 38; DB 1; Length 513;
Best Local Similarity 53.8%; Pred. NO. 48;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Search completed: July 7, 2003, 14:21:59
Job time : 1.45283 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:44 ; Search time 1.63019 Seconds
(without alignments)
2022.314 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78

Sequence: 1 MRYRASALGSDGVRV 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	1044	11	Q9ERB6
2	50	64.1	260	16	Q9ETD6 mus musculus
3	49	62.8	262	16	Q9CP73 pasteurella
4	48	61.5	99	2	Q827C7
5	45	57.7	264	16	Q8Y072
6	44	56.4	268	2	Q8W034
7	43	55.1	826	2	Q8R033
8	42	53.8	291	10	Q9F1K5
9	41	52.6	151	2	Q9LPE0
10	41	52.6	151	2	Q9LPE6
11	41	52.6	151	16	Q935M8
12	41	52.6	252	16	Q92M00
13	41	52.6	258	16	Q930L5
14	41	52.6	356	10	Q81792
15	41	52.6	415	10	Q9M068
16	41	52.6	419	10	Q8VZ07

17	41	52.6	572	16	Q9S1S8	Q9S1S8 streptomyc
18	40	51.3	375	16	Q9F2K4	Q9F2K4 streptomyc
19	40	51.3	412	4	Q96PW2	Q96PW2 homo sapien
20	40	51.3	437	16	Q9KYZ2	Q9KYZ2 streptomyc
21	40	51.3	687	5	Q9GXY9	Q9GXY9 armigeres s
22	39	50.0	242	2	Q9L8E7	Q9L8E7 vibrio harv
23	39	50.0	261	16	Q9K1S1	Q9K1S1 neisseria m
24	39	50.0	261	16	Q9JSS8	Q9JSS8 neisseria m
25	39	50.0	262	16	Q92D51	Q92D51 listeria in
26	39	50.0	262	16	Q8Y8D5	Q8Y8D5 listeria in
27	39	50.0	274	16	Q8Y0D6	Q8Y0D6 listeria mo
28	39	50.0	467	10	Q8S1Q8	Q8S1Q8 oryza sativ
29	39	50.0	551	2	Q8R233	Q8R233 anabaena sp
30	39	50.0	793	2	Q94857	Q94857 streptomyc
31	39	50.0	793	2	Q94857	Q94857 streptomyc
32	39	50.0	844	10	Q94BN8	Q94BN8 lactobacill
33	39	50.0	875	10	Q9FLQ1	Q9FLQ1 arabidopsis
34	39	50.0	924	16	P73172	P73172 synecocyst
35	39	50.0	1296	16	Q9P9W1	Q9P9W1 xylella fas
36	39	50.0	1470	5	Q9G193	Q9G193 leishmania
37	39	50.0	2393	2	Q925F4	Q925F4 myxococcus
38	38.5	49.4	513	16	Q8X519	Q8X519 escherichia
39	38	48.7	110	16	Q92SV3	Q92SV3 rhizobium m
40	38	48.7	146	10	Q64874	Q64874 arabidopsis
41	38	48.7	192	5	Q8T592	Q8T592 giardia lam
42	38	48.7	254	16	Q92PP8	Q92PP8 rhizobium m
43	38	48.7	272	16	Q8UHC5	Q8UHC5 agrobacteri
44	38	48.7	298	10	Q9FJ82	Q9FJ82 arabidopsis
45	38	48.7	312	16	Q97Q18	Q97Q18 streptococc

ALIGNMENTS

RESULT 1	ID	Q9ERB6	PRELIMINARY:	PRT:	1044 AA.
AC	Q9ERB6	01-MAR-2001 (TREMBL)	16	Created	
DT	01-MAR-2001 (TREMBL)	16	Last sequence update)		
DT	01-JUN-2002 (TREMBL)	21	Last annotation update)		
DE	Nucleic myosin I beta.				
GN	MYOIC.				
OS	Mus musculus (Mouse).				
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Pestic-Dragovich L., Stojiljkovic L., Phillimonenko A.A., Nowak G.,				
RA	Ke Y., Settlage R.E., Shabanowitz J., Hunt D.F., Hozak P.,				
RA	"A Myosin I Isoform in the Nucleus.";				
RT	Science 0:0-0(2000).				
RL	EMBL; AY007255; AAC02570.1; -				
DR	HSSP; P08799; 1MND.				
DR	MCD; MGI:106612; MyoIC.				
DR	InterPro; IPR000048; IO.region.				
DR	InterPro; IPR001609; myosin_head.				
DR	Pfam; PF00063; myosin_head; 1.				
DR	PRINTS; PR00193; MYOSINHEAVY.				
DR	PRODOM; PD000355; myosin_head; 1.				
DR	SMART; SM00015; IQ; 2.				
DR	SMART; SM00242; MYSC; 1.				
SO	SEQUENCE	1044 AA;	119876 MW;	CDBOEF699D041C9 CRC64;	
Query Match					
Best Local Similarity 100.0%; Score 78; DB 11; Length 1044;					
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY 1 MRYRASALGSDGVRV 16					

Db 1 MRYRASALGSDGVRV 16

RESULT 2

Q9CP73 PRELIMINARY; PRT; 260 AA.
AC Q9CP73;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Fadi.
GN FADI OR PM0182.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RX May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1 SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL: AE006052; AK02266.1; -
DR HSSP: P29132; IQSG.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 260 AA; 27850 MW; 922E5B8B0C8963 CRC64;

Query Match

Best Local Similarity 64.1%; Score 50; DB 16; Length 260;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 MRYRASALGSDGVRV 15
Db 171 MRYRASALGSDGVRV 184

RESULT 3

0827C7 PRELIMINARY; PRT; 262 AA.
AC 0827C7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Enoyl-[acyl-carrier-protein] reductase (NADH).
GN STY1352.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RX Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Peltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RT Nature 413:848-852(2001).
RL EMBL: AL627270; CAD01621.1; -
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
KW Complete proteome.

SQ SEQUENCE 262 AA; 27775 MW; 1844E8792E709A93 CRC64;

Query Match
Best Local Similarity 62.8%; Score 49; DB 16; Length 262;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVRV 15
Db 170 MRYRASALGSDGVRV 184

RESULT 4

Q51891 PRELIMINARY; PRT; 99 AA.
AC Q51891;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to E.coli EnvM (Fragment).
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=584;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=AF 2924;
RX MEDLINE=93170270; PubMed=8436105;
RX Magnogna G., Allicati N., Aceto A., Piccolomini R., Di Ilio C.,
RA Barra D., Martini F.;
RT "The amino acid sequence of glutathione transferase from Proteus
RT mirabilis, a prototype of a new class of enzymes.";
RT Eur. J. Biochem. 211:421-425(1993).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=AF 2924;
RX MEDLINE=9636500; PubMed=8761466;
RX Perito B., Allicati N., Aceto A., Casalone E., Di Ilio C., Masulli M.,
RA Dragani B., Polistelli M.;
RT "Molecular cloning and overexpression of a glutathione transferase
RT gene from Proteus mirabilis.";
RL Biochem. J. 318:157-162(1996).
DR EMBL: U38482; AAC44361.1; -
DR HSSP: P29132; IDPI.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
FT NON_TER
SQ SEQUENCE 99 AA; 10395 MW; 2E442AB04AFA031 CRC64;

Query Match
Best Local Similarity 61.5%; Score 48; DB 2; Length 99;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVRV 15
Db 7 MRYRASALGSDGVRV 21

RESULT 5

Q8Y072 PRELIMINARY; PRT; 264 AA.
AC Q8Y072;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Probable enoyl-[acyl-carrier-protein] reductase (EC 1.3.1.9).
GN FADI OR RSC1172 OR RS04528.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=GM1100;

RX MEDLINE=21681879; PubMed=11823852;
 RA Salenobat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chendrier M., Choise N., Claudel R., Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Siglier P., Thebaud P., Whalen M., Winkler P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT Genome sequence of the plant pathogen *Ralstonia solanacearum*.
 RL Nature 415:497-502(2002).
 DR EMBL: AL646063; CAD14874.1;
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 264 AA; 28006 MW; 1CE9DB4C10AE3A52 CRC64;

Query Match 57.7%; Score 45; DB 16; Length 264;
 Best Local Similarity 64.3%; Pred. No. 13;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 14
 DB 170 VRYLASALGPKGR 183

RESULT 6

ID 08WQ34 PRELIMINARY; PRT; 444 AA.

AC 08WQ34;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical predicted transmembrane protein P265.17, unknown
 DE function.
 GN P265.17.
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Robben J., Grymonprez B., Waijens I., Aert R., Volckaert G.,
 RA Ivens A.C., Quail M., Rajendram M.A., Barrell B.G.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome."
 RL Genome Res. 8:135-145(1998).
 DR EMBL: AL359716; CAD19425.1; -.
 KW Transmembrane.
 SQ SEQUENCE 444 AA; 49507 MW; 5E0CBA3C2868398A CRC64;

Query Match 57.7%; Score 45; DB 5; Length 444;
 Best Local Similarity 56.2%; Pred. No. 22;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 16
 DB 22 MYSASLSSSGIEVT 37

RESULT 7

ID 09RE03 PRELIMINARY; PRT; 268 AA.

AC 09RE03;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Enoyl reductase.
 GN FABL.

OS Legionella pneumophila.
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group.
 OC Legionellaceae; Legionella.
 NCBI_TaxID=446;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PHILADELPHIA 1 SGI;
 RA Kabbash C.A., Albala A., Latza P.D., Feinmark S., Weiss I., Walthe E.,
 RA Silverstein S.C., Shuman H.A.;
 RT "Antibiotic activity of gemfibrozil for *L. pneumophila* and
 RT *M. tuberculosis*."
 RL Submitted (MUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC EMBL: Y14405; CAB65183.1; -.
 DR HSSP: P29132; 10SG.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Oxidoreductase.

SQ SEQUENCE 268 AA; 28466 MW; DF6B4E16AE589E2D CRC64;

Query Match 56.4%; Score 44; DB 2; Length 268;
 Best Local Similarity 53.3%; Pred. No. 19;
 Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 15
 DB 177 VRYLASLGRGRL 191

RESULT 8

ID 09FLK5 PRELIMINARY; PRT; 826 AA.

AC 09FLK5;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Chitinase.
 GN BCC1.
 OS Burkholderia cepacia (Pseudomonas cepacia).
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
 OC Burkholderia.
 NCBI_TaxID=292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matsuda Y., Kakutani K., Nonomura T., Ikeda S., Tamai T., Kuwabara Y.,
 RA Tanaka H., Toyoda H.;
 RT "DNA sequence of chitinase gene cloned from *Burkholderia cepacia*."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB053088; BAB20043.1; -.
 DR HSSP: O54468; 10BA.
 DR InterPro: IPR004865; CarB_Hex.
 DR InterPro: IPR001540; GH_20.
 DR Pfam: PF03173; CHB_Hex; 1.
 DR Pfam: PF00728; Glyco_Hydro_20; 1.
 DR PRINTS: PR00738; GLHYDRASE20.
 SQ SEQUENCE 826 AA; 90168 MW; E9CC7A6544DC31EC CRC64;

Query Match 55.1%; Score 43; DB 2; Length 826;
 Best Local Similarity 60.0%; Pred. No. 97;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 15
 DB 252 LRDRATTLGIDGARV 266

RESULT 9

ID 09LPK0 PRELIMINARY; PRT; 291 AA.

AC 09LPK0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

```

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F6N18.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids 11; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Chao C., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Alatafi H., Bel O., Chin C., Chiu J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thavert A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F6N18 from chromosome
RT 1."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC017118; AAF25981.1; -.
SQ SEQUENCE 291 AA; 32361 MW; 3F64694A4C8397EC CRC64;

Query Match 53.8%; Score 42; DB 10; Length 291;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGV 12
DB 130 MRYRASALGSDG 141
|||||
|

RESULT 10
ID Q9LSF6 PRELIMINARY; PRT; 151 AA.
AC Q9LSF6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Orf, hypothetical protein.
GN R0189.
OS Salmonella typhi.
OG Plasmid R27.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20280091; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grobeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer."
RL Nucleic Acids Res. 28:2177-2186(2000).
DR EMBL; AF250878; AAF70026.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 151 AA; 18075 MW; 15E31FA26518CEB9 CRC64;

Query Match 52.6%; Score 41; DB 2; Length 151;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 RYRASALGSDGV 15
DB 107 RYRLNRIGTDGLTV 120
|||||
|

RESULT 11
ID Q935M8 PRELIMINARY; PRT; 151 AA.
AC Q935M8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

```

```

DE Putative membrane protein.
GN HCM1.128.
OS Salmonella typhi.
OG Plasmid pHC1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Kirogh A., Larsen T.S., Leather S., Moule S., O'Gea P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AL513383; CAD09726.1; -.
KW Plasmid; Complete proteome.
SQ SEQUENCE 151 AA; 18158 MW; A9927F14B3C3A30F CRC64;

Query Match 52.6%; Score 41; DB 16; Length 151;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 RYRASALGSDGV 15
DB 107 RYRLNRIGTDGLTV 120
|||||
|

RESULT 12
ID Q92M00 PRELIMINARY; PRT; 252 AA.
AC Q92M00;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative oxidoreductase protein (EC 1.1.1.-).
GN R02563 OR SMC02336.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=1021;
RA MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purrelle B., Ransperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47142.1; -.
DR InterPro: IPR002198; ADL-short.
DR Pfam: PF00106; adh_short.1.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN.1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 252 AA; 26560 MW; 0B10B7A0DBEF6628 CRC64;

Query Match 52.6%; Score 41; DB 16; Length 252;
Best Local Similarity 57.1%; Pred. No. 57;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGV 15
|||||
|

```


DB 167 RHLAHAFPGDGIHV 180

RESULT 13

0930L5 PRELIMINARY; PRT; 258 AA.

AC 0930L5: 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Putative.

GN RAO180 OR SMA0335.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Plasmid pSymba (megaplasmid 1).

CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

NCBI_TaxID=382;

RM [1]

RP SEQUENCE FROM N.A.

RC STRAIN-1021.

RX MEDLINE-21396509; Pubmed-11481432;

RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,

RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,

RA Gurjal M., Hong A., Hultzer L., Hyman R.W., Kahn D., Kahn M.E.,

RA Kallman S., Keating D.H., Palm C., Peck M.C., Stuzky R., Wells D.H.,

RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;

RT "Nucleotide sequence and predicted functions of the entire

RT Sinorhizobium meliloti pSymba megaplasmid."

RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).

DR EMBL: AEO07211; AAK64838.1; -;

DR InterPro: IPR002198; ADH_short.

DR Pfam: PF00106; adh_short.1.

DR PROSITE: PS00061; ADH_SHORT; UNKNOWN.1.

KW Plasmid; Hypothetical protein; Complete proteome.

SQ SEQUENCE 258 AA; 26146 MW; 446BABC14A2AB8B1 CRC64;

Query Match 52.6%; Score 41; DB 16; Length 258;

Best Local Similarity 72.7%; Pred. No. 59;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 ASALGSDGVRV 15

DB 177 ADALGPDGIHV 187

RESULT 14

081792 PRELIMINARY; PRT; 356 AA.

AC 081792: 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Protein kinase - like protein (Fragment).

GN F8D20.110.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC euroside II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;

RM [1]

RP SEQUENCE FROM N.A.

RA Koester P., Hempel S., Entian K.-D., Hohnsels J., Jesse T.,

RA Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;

RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AL031135; CAA20030.1; -;

DR InterPro: IPR000719; Euk_kinase.

DR InterPro: IPR002290; Ser_thr_kinase.

DR InterPro: IPR004040; STY_kinase.

DR InterPro: IPR001245; Tyr_kinase.

DR Pfam: PF00069; kinase.1.

DR PRINTS: PR00109; TYRKINASE.

DR ProDom: PD000001; Euk_kinase.1.

DR SMART: SM00221; STYK: 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

KW Kinase; Serine/threonine-protein kinase.

FT NON_TER 1

SQ SEQUENCE 356 AA; 40012 MW; FAD9736384FB8221 CRC64;

Query Match 52.6%; Score 41; DB 10; Length 356;

Best Local Similarity 53.3%; Pred. No. 84;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RYRASALGSDGVRV 16

DB 333 RYKAGALGAEKRRAT 347

RESULT 15

09M068 PRELIMINARY; PRT; 415 AA.

AC 09M068: 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Protein kinase-like protein.

GN ATG35600.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC euroside II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;

RM [1]

RP SEQUENCE FROM N.A.

RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;

RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AL161587; CAB80276.1; -;

DR InterPro: IPR000719; Euk_kinase.

DR InterPro: IPR002290; Ser_thr_kinase.

DR InterPro: IPR004040; STY_kinase.

DR InterPro: IPR001245; Tyr_kinase.

DR Pfam: PF00069; kinase.1.

DR PRINTS: PR00109; TYRKINASE.

DR ProDom: PD000001; Euk_kinase.1.

DR SMART: SM00221; STYK: 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 415 AA; 45879 MW; 2FAFD025A219F528 CRC64;

Query Match 52.6%; Score 41; DB 10; Length 415;

Best Local Similarity 53.3%; Pred. No. 99;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RYRASALGSDGVRV 16

DB 392 RYKAGALGAEKRRAT 406

Search completed: July 7, 2003, 14:23:58

Job time: 4.63019 secs

